PRECISION COSMOLOGY WITH WEAK GRAVITATIONAL LENSING

by

Andrew P. Hearin

B.S. in Physics & Mathematics, University of Richmond, 2003

M.S. in Mathematics, University of Massachusetts at Amherst, 2006

MSc in Theoretical Physics, Imperial College of London, 2007

Submitted to the Graduate Faculty of

the Department of Physics and Astronomy in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF PHYSICS AND ASTRONOMY

This dissertation was presented

by

Andrew P. Hearin

It was defended on

20 July, 2012

and approved by

Andrew Zentner, University of Pittsburgh

Jeff Newman, University of Pittsburgh

Michael Wood-Vasey, University of Pittsburgh

Vittorio Paolone, University of Pittsburgh

Rupert Croft, Carnegie Mellon University

Dissertation Director: Andrew Zentner, University of Pittsburgh

PRECISION COSMOLOGY WITH WEAK GRAVITATIONAL LENSING

Andrew P. Hearin, PhD

University of Pittsburgh, 2012

In recent years, cosmological science has developed a highly predictive model for the universe on large scales that is in quantitative agreement with a wide range of astronomical observations. While the number and diversity of successes of this model provide great confidence that our general picture of cosmology is correct, numerous puzzles remain. In this dissertation, I analyze the potential of planned and near future galaxy surveys to provide new understanding of several unanswered questions in cosmology, and address some of the leading challenges to this observational program. In particular, I study an emerging technique called cosmic shear, the weak gravitational lensing produced by large scale structure. I focus on developing strategies to optimally use the cosmic shear signal observed in galaxy imaging surveys to uncover the physics of dark energy and the early universe.

In chapter 1 I give an overview of a few unsolved mysteries in cosmology and I motivate weak lensing as a cosmological probe. I discuss the use of weak lensing as a test of general relativity in chapter 2 and assess the threat to such tests presented by our uncertainty in the physics of galaxy formation. Interpreting the cosmic shear signal requires knowledge of the redshift distribution of the lensed galaxies. This redshift distribution will be significantly uncertain since it must be determined photometrically. In chapter 3 I investigate the influence of photometric redshift errors on our ability to constrain dark energy models with weak lensing. The ability to study dark energy with cosmic shear is also limited by the imprecision in our understanding of the physics of gravitational collapse. In chapter 4 I present the stringent calibration requirements on this source of uncertainty. I study the potential of weak lensing to resolve a debate over a long-standing anomaly in CMB measurements in chapter

5. Finally, in chapter 6 I summarize my findings and conclude with a brief discussion of my outlook on the future of weak lensing studies of cosmology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to begin by expressing my gratitude to my advisor, Andrew Zentner, whose contribution to my development as a scientist is singular. His unfailingly patient guidance at each stage of my graduate career has been invaluable. I cannot thank him enough. I also thank the other four members of my dissertation committee for their feedback on my work throughout my time at Pitt. In addition to his feedback, Rupert Croft has been accommodating beyond the call of duty in the coordination of my committee meetings, which is worthy of special recognition. I am indebted to Vittorio Paolone for the interest he has taken in my work and for his patience with my bald attempts at bribery with mixes of obscure 60's garage rock recordings. My thanks to Michael Wood-Vasey for consistently asking thoughtful, insightful questions and for demonstrating by way of example how to maintain a broad intellectual curiosity while remaining focused on productive scientific work. I give special thanks to Jeff Newman, whose generosity with his time in all of my years at Pitt has been truly remarkable. He has had an immeasurable impact on both my technical skills as well as my general scientific worldview, for which I will always be grateful.

I thank my graduate student friends who shared Allen 419 with me, particularly Brian Cherinka and Dan Matthews for indulging my procrastinatory habit of posing questions concerning their preference between preposterous hypothetical scenarios with no desirable outcome.

I thank my co-authors from other institutions, Zhaoming Ma, Dragan Huterer, Cameron Gibelyou, and Andreas Berlind. In addition to being grateful for their scientific contributions to our publications together, they have each of them come to know my propensity for email verbosity and have been nothing but patient.

My thanks are also due to numerous professors who have contributed to my education

before my arrival in Pittsburgh. At Imperial College, Carlo Contaldi played a critical role in guiding me towards a degree in cosmology, and Fay Dowker's clarity of thinking continues to serve as a gold standard example for my own. Ted Bunn at the University of Richmond instilled in me a drive for geometric intuition and continues to serve as an intellectual role model. Pierson Marshall at Isidore Newman is perhaps most responsible for awakening my scientific curiosity; his influence on my academic pursuits has been profound, and I express my enduring gratitude to him for it.

I thank my family for their continued support of my education. In particular, I thank Nathan and Lauren for taking an interest in my work and for enabling my undisguised agenda to convince their son, Jonathan, that science and reading are cool. I thank my brother for his steady supply of encouragement throughout my graduate career, and my mother for her unfailing support of whatever path I have chosen. I thank my dad, without whose transferred trait of relentlessness this thesis would not have been possible. Finally, I thank my partner, Robin, for sharing in my triumphs and disappointments, large and small.

1.0 OVERVIEW

1.1 INSTRUCTIONS FOR READERS

This thesis collects together a variety of results pertaining to the study of cosmology with weak lensing, and its chapters are relatively self-contained. This structure allows for readers with different backgrounds to focus on particular chapters or sections according to their specialized interests. Readers seeking a broad, relatively non-technical overview of the science I study in this dissertation may wish to focus primarily chapter 1 . Weak lensing scientists with a particular interest in photometric redshift errors will likely find chapter 3 the most relevant, while those specializing in nonlinear structure growth might wish to concentrate on chapters 2 and 4. Readers interested in the possibility of weak lensing opening up a new window onto a persistent large-angle CMB anomaly may skip directly to chapter 5. Those who are already well-versed in weak lensing may find the bullet-point style Summary sections concluding each chapter sufficient to quickly review my main results.

1.2 BACKGROUND AND MOTIVATION

Cosmology is the study of the physics of the universe on the largest scales. In the early 20^{th} century, revolutionary developments in both astronomy and theoretical physics established a firm scientific foundation for the study of cosmology. Progress through the middle of the 20^{th} century was sporadic and comparatively slow as extragalactic astronomy focused more on targeted observations of relatively small collections of objects rather than large-scale surveys. Advances in digital detection equipment and computation over the last few decades have fa-

cilitated the arrival of so-called precision cosmology, a moniker that reflects the character of the questions pursued by contemporary cosmologists. Numerous observational programs today are producing vast, rich data sets that are being mined to provide percent-level constraints on a small set of parameters describing our model for the history of the universe and its contents. In particular, current and near future galaxy surveys hold great promise for answering a number of outstanding questions in fundamental physics. In this dissertation, I study the physics of a variety of cosmological signals that will be seen by these surveys and investigate several of the leading challenges that will need to be met in order to realize the potential of these observational efforts.

In the modern cosmological model, generally referred to as ΛCDM, on large scales matter and energy are homogeneously distributed in space and the universe has been expanding throughout its 13.7 billion year lifetime. The first astronomical evidence providing quantitative support for this model came nearly one hundred years ago, when galaxies in all directions were observed to be receding from us. The recession velocity of a galaxy is typically quantified by its *redshift*, the fractional shift in the wavelength of light emitted by the galaxy. Through observations made in the early 20^{th} century by Vesto Slipher, Edwin Hubble, and others it was determined that the redshift of a galaxy increases with our distance to it, a generic consequence of a cosmological model in which space itself is homogeneously expanding.

Since this discovery, our understanding of the expanding universe has been refined into a very specific, predictive model that has proven to be quantitatively consistent with a broad range of independent astronomical measurements. One of the most compelling pieces of evidence supporting this model came with the discovery of the Cosmic Microwave Background (CMB). Since the ΛCDM model predicts that space has been expanding throughout its history, if we wind back the cosmic clock far enough there must have been a time in which the universe was very hot and densely packed with its homogeneously distributed energy contents. Such a physical system would produce a blackbody radiation field uniformly distributed throughout all of space. In 1964, Arno Penzias and Robert Woodrow Wilson presented the first detection of this microwave radiation field, earning them the Nobel Prize in physics in 1978 for their discovery. Subsequent observations of the CMB revealed that it

exhibits a blackbody spectrum purer than has ever been produced in a terrestrial laboratory, providing strong support for our contemporary cosmological model.

While ΛCDM has achieved great success in predicting astronomical observations, numerous puzzles remain. For example, roughly 5/6 of the gravitating matter in the universe is composed of dark matter particles that emit no light and have thus far evaded all attempts to study them with particle physics equipment. Naturally, this makes dark matter particles difficult to study. Evidence for the existence of this dark matter is myriad, and includes the rotation curves of galaxies, the primordial abundance of the light atomic elements, and various measurements of the growth of cosmic structure. These observations constitute detections of the gravitational influence of dark matter, but the nature of the non-gravitational interactions enjoyed by the dark matter particle(s) remains unknown.

One of the most mysterious components of our cosmological model goes by the name of *dark energy*. In 1999, through observations of Type Ia supernovae two collaborations of scientists determined that the expansion of the universe is accelerating. The universally attractive nature of gravity implies that the presence of matter in an expanding universe should only slow down the rate of expansion, and so this acceleration points to a fundamental gap in our understanding of the universe. The cause of the accelerating expansion has been dubbed dark energy. Indeed dark energy is simply a name for the fundamental puzzle that is the nature of cosmic acceleration and an enormous amount of effort is now being devoted to determine the causative agent of the cosmic acceleration.

One newly emerged technique for studying the physics of dark energy is weak lensing. Einstein's theory of General Relativity predicts that spacetime is curved by the matter inhabiting it. One consequence of this prediction is that photons are deflected as they pass by clumps of matter, causing the path of light rays in the presence of matter to be bent relative to the path traversed by light traveling through empty space. This phenomenon is referred to as gravitational lensing. When there is a sufficiently large clump of mass along the line of sight from our telescopes to a distant source of light, the curved path of photons around the intervening matter causes us to infer that the light was emitted from a different position on the sky than the true position. The deflection angle is a function both of the amount of mass responsible for the lensing as well as the relative distance between the

observer, the massive lens, and the source. Measurements of the deflection of the positions of stars by the Sun during a total solar eclipse in 1919 provided one of the first experimental confirmations of relativity.

When a source of light has spatial extent (in other words, when it is not a point source), gravitational lensing can change both the apparent position and shape of the source. Since gravitational lensing is sensitive to the mass of the lens, measuring the shape distortions of light sources behind a distant clump of matter allows one to measure the mass of the foreground clump. Indeed, exploiting shape measurements of galaxies behind a galaxy cluster is one of the most widely used techniques in contemporary astronomy to measure cluster mass. One of the primary reasons that gravitational lensing has become such a ubiquitous mass determination method is that no assumptions about the type of particles that make up the massive lens are required: both dark matter and ordinary matter produce the same deflection of light per unit mass because in general relativity spacetime distortion is determined strictly by mass.

In this dissertation, I will focus on a manifestation of gravitational lensing known as *cosmic shear*, a term referring to the distortion of the images of galaxies produced by large-scale inhomogeneities in the distribution of matter throughout the universe[1]. In contrast to lensing by galaxy clusters, in cosmic shear there is no single, distinct lens that can be associated with an individual deflection event. Rather, the cosmic shear signal refers to the aggregate effect of numerous, tiny deflection events produced by all of the large scale structure distributed along the line of sight to the source of light. Thus the cosmic shear signal is stochastic: the strength and character of the signal reflect the statistical distribution of matter rather than the mass of, or distance to, any particular lens. The cosmic shear signal was first detected over a decade ago, but as we will see, measurements of dark energy properties with cosmic shear require observations of hundreds of millions of galaxies, and so it will not be possible to use this signal to provide constraints on dark energy that are competitive with other techniques until galaxy imaging surveys such as the Subaru Hyper-Suprime Cam and the Dark Energy Survey begin taking data later this year.

[1]Since I do not study any form of weak lensing other than cosmic shear throughout this dissertation, I will use the terms "weak lensing" and "cosmic shear" interchangeably as there should be no confusion over the physical scenario to which I am referring.

As mentioned above, gravitational lensing depends both on the mass of the lens and the relative distance between the observer, lens, and source. Because of the stochastic nature of cosmic shear, these dependencies take on a different character than in the conceptually simpler case in which there is a single massive lens. In cosmic shear, the mass dependence of the signal figures in as the statistical distribution of over- and under-dense regions, quantified by the matter power spectrum, $P_\delta(k)$; the signal is dependent on distance via its dependence on the expansion history of the universe. This basic physical picture allows us to understand how measurements of weak lensing by large scale structure can be exploited to test theories of dark energy. Different models of the cosmic acceleration make different predictions for both the growth of structure and for the expansion history. Since cosmic shear is sensitive to both structure growth and the expansion history, measurements of the cosmic shear signal allow for the possibility of distinguishing between competing models. Alternatively, the cosmic shear signal can be used to test the validity of general relativity. For a fixed set of cosmological parameters, general relativity makes a unique prediction for the expansion history of the universe and for the growth of cosmic structure. Thus weak lensing measurements can be used to test the consistency of the observed structure growth and expansion history with one another within the GR framework.

After reviewing the general phenomenology of weak lensing tests of general relativity on cosmological scales, I begin chapter 2 by addressing a major challenge to this program. Interpreting the weak lensing signal requires a detailed understanding of the nonlinear physics of gravitational collapse. It is possible to excise cosmic shear data on the relatively small scales (a few Megaparsecs) where nonlinear effects are the most dramatic, but such an excision significantly degrades dark energy constraints. Modeling nonlinear structure formation will thus be essential in order for forthcoming galaxy imaging surveys to realize their potential for constraining dark energy and alternative models of gravity.

The problem of making precise predictions for the distribution of matter on Mpc scales is further complicated by the significant fraction ($\sim 1/6$) of the mass in the universe that is comprised of baryons[2]. Significant rearrangement of dark matter may result from baryonic

[2]By the term *baryons*, in this context, I am simply referring to Standard Model particles as opposed to dark matter particles, the large scale behavior of the latter being dominated by gravity-only interactions.

processes at work during the formation and evolution of galaxies. For example, complex astrophysical phenomena such as the radiative cooling of intergalactic gas and active galactic nuclei have an important influence on the distribution of matter on scales that are important for weak lensing. The uncertainty associated with modeling these and other baryonic processes results in a concomitant uncertainty in the interpretation of weak lensing observables, manifesting as modified structure growth for a fixed cosmic distance scale. Using two proposed methods, in chapter 2 I show that one could be led to reject the null hypothesis of general relativity when it is the true theory if uncertainty in baryonic processes is neglected. Recent simulations suggest that we can correct for baryonic effects using a parameterized model describing the concentration of the total mass contained in the dark matter halos in which galaxies are embedded. Motivated by these findings, I study the ability of future weak lensing surveys to constrain the internal structures of halos and test the null hypothesis of general relativity simultaneously.

Another challenge to weak lensing studies of the cosmic acceleration arises from the need to determine the redshifts to the lensed, source galaxies using photometric, rather than spectroscopic, information. This necessity is driven by the "weakness" of the weak lensing signal: in cosmic shear, a typical gravitational lensing event only distorts the shape of a background galaxy image by $\sim 1\%$, but the intrinsic variance in the elliptical shapes of galaxies is $\sim 30\%$, and so photometric images of hundreds of millions of galaxies are required in order to attain a signal-to-noise ratio that is sufficient for precision studies of cosmology. Spectroscopic measurements of this many galaxies would require far too much observing time to be practical, and so forthcoming lensing studies must rely on photometrically-determined redshifts of the lensed galaxies.

Photometric redshifts, hereafter *photo-z's*, are intrinsically less precise than spectroscopic redshifts since a tremendous amount of information contained in the full spectral energy distribution of a galaxy is washed out by only observing the galaxy through a few broadband filters. This imprecision presents one of the primary challenges to weak lensing studies of dark energy. In chapter 3 I provide a detailed account of the influence of photo-z errors on the dark energy error budget. I explore the relative importance of calibrating the photo-z's of a core set of relatively well-understood galaxies as compared to the need to identify

potential outliers in the redshift distribution of the lensed galaxies. In so doing, I provide an approximate blueprint for computing the influence of specific photo-z outlier populations on dark energy constraints. This work both enables an assessment of the success of photometric redshift determination techniques in terms of dark energy parameter estimation, as well as provides new strategies to optimize weak lensing data analyses.

As discussed in the Appendix, the degradation of the constraining power of weak lensing with photo-z uncertainty differs depending on how one models the physics of gravitational collapse. These results demonstrate that there is a nontrivial interplay between the nonlinear growth of cosmic structure and the redshift distribution of the sources used to measure the weak lensing signal. This suggests the need to model uncertainty in photo-z's and structure growth in conjunction with one another. In chapter 4 I present results concerning the stringent calibration requirements for weak lensing analyses of future imaging surveys that simultaneously addresses photo-z uncertainty and errors in the calibration of the matter power spectrum. I find that when photo-z uncertainty is taken into account, the requirements on the level of precision in the prediction for the matter power spectrum are more stringent than previously thought. I compare these updated calibration requirements to the contemporary state-of-the-art in photometric redshift estimation and predictions of the power spectrum and suggest strategies to utilize forthcoming data optimally.

In chapter 5 I shift attention from dark energy to another aspect of ΛCDM with lingering questions: inflation. The very early universe is thought to have experienced an inflationary phase when space expanded extremely rapidly. The concept of inflation was originally introduced to solve several major problems in modern cosmology, including why our universe has flat spatial geometry and why magnetic monopoles are not observed in vast abundance. One attractive feature of inflation is that the very mechanism employed to solve these problems presents a natural candidate for the production of the tiny initial seeds of over- and underdensities that eventually grew to form the stars, galaxies, and clusters that populate the universe today. Like all scientific advancements, the development of the theory of inflation gave rise to a new set of questions (for example, by what mechanism did the scalar field responsible for the inflationary expansion decay into standard model particles?), many of which scientists still pursue today.

One of the most powerful ways to constrain models of inflation comes from observations of the CMB. As discussed above, the CMB is very nearly homogeneous, but inhomogeneities at a level of one part in 10^5 have been measured with very high precision. In ΛCDM, these inhomogeneities are the seeds from which cosmic structure forms and are thought to have originated from fluctuations in the inflaton field that drove the inflationary epoch. One of the most elementary statistics that can be constructed to describe the observed fluctuations about the mean temperature is the angular correlation function, $\mathcal{C}(\theta)$, which quantifies how hot and cold spots in the CMB are correlated with one another as a function of their angular separation θ. The near-total lack of correlation in the temperature distribution is a long-standing anomaly in CMB science: $\mathcal{C}(\theta)$ is observed to be near zero for angular scales larger than 60 degrees. It remains unclear whether this large-angle anomaly is simply a statistical fluke or a signal of exotic physics that took place during the inflationary epoch, but in the case of the latter the signature of this physics should in principle also be imprinted in the distribution of large-scale structure from which the CMB fluctuations evolved. In chapter 5 I investigate the potential for future galaxy imaging surveys to test the authenticity of the large-angle CMB temperature correlation deficit, thereby yielding new insight into the physics of the early universe. I conclude this dissertation in chapter 6 with a brief, semi-technical summary of my primary findings and a discussion of my outlook for the future of ΛCDM studies with weak lensing.

1.3 PUBLICATION LIST

The work presented in Chapters 2-5 is primarily based on four publications completed through the course of my graduate work. For convenience, I list the references to those publications here:

- A.P. Hearin, A.R. Zentner, *The Influence of Galaxy Formation Physics on Weak Lensing Tests of General Relativity,* JCAP, 4, 32 (2009), ArXiv:0904.3334
- A.P. Hearin, A.R. Zentner, Z. Ma, D. Huterer, *A General Study of the Influence of Catastrophic Photometric Redshift Errors on Cosmology with Cosmic Shear Tomography,*

ApJ, 720, 1351 (2010), ArXiv:1002.3383

- A.P. Hearin, A.R. Zentner, Z. Ma, *General Requirements on Matter Power Spectrum Predictions for Cosmology with Weak Lensing Tomography*, JCAP, 4, 34, (2011), ArXiv:1111.0052
- A.P. Hearin, C. Gibelyou, A.R. Zentner, *Testing the Origin of the CMB Large-Angle Correlation Deficit with a Galaxy Imaging Survey*, JCAP, 10, 012 (2011), ArXiv:1108.2269

2.0 TESTING GENERAL RELATIVITY WITH COSMIC SHEAR

2.1 INTRODUCTION

Current data indicate that the cosmological expansion is accelerating [6; 50; 144; 152; 153; 174; 182; 183; 193]. In broadest terms, two options have been put forth to describe accelerated cosmic expansion. The first is that cosmic acceleration is caused by some as yet unidentified contribution to the stress energy of the universe. This option includes vacuum energy (observationally indistinguishable from a cosmological constant) and dynamical models of dark energy. The second option is that gravity deviates from the general relativistic description on large scales, an option we refer to as modified gravity.

Viable alternative options to general relativity for which definite predictions have been made are few and far between, so many authors have suggested that a fruitful way to apply forthcoming data will be to check for the mutual consistency of different observable phenomena with the predictions of general relativity [79; 111; 112; 127; 189; 200; 201; 202]. The basic idea is that one may obtain observational handles on both the distance-redshift relation and the growth rate of cosmic expansion. Within general relativity, both of these can be predicted from the same dark energy parameters, so it is possible to measure, for example, distance and then check for consistency with the growth of cosmic structure. Forthcoming weak gravitational lensing surveys will provide the most powerful means to probe the matter distribution in the universe directly and are an indispensable piece of any such consistency check [79; 202]. A significant literature already exists detailing the power of tomographic weak lensing to constrain dark energy parameters, including Refs. [4; 14; 43; 65; 72; 76; 78; 148; 149; 173; 179; 180; 197; 198]. In fact, the effectiveness of any given matter fluctuation to serve as a lens is sensitive to the distance scale of the universe and so weak gravitational

lensing alone will provide a very powerful check for the consistency of general relativity. In this chapter, we study the effectiveness of such a consistency check in light of recent studies that indicate that predictions for weak lensing observables will contain significant inherent uncertainties due to the poorly-understood behavior of the baryonic component of the universe [89; 155; 191; 197; 199].

The strategy for checking the consistency of general relativity and searching for a sign of modified gravity outlined in the previous paragraph seems simple enough. However, making predictions for weak lensing power spectra is fraught with numerous practical difficulties. Weak lensing will use information on scales where density fluctuations are well beyond the linear regime (relative overdensities $\delta \gtrsim 1$). To utilize this information at the level that forthcoming observational programs will permit, numerical simulations must be able to predict the nonlinear matter power spectra to better than a percent [80; 82]. Sufficient precision should be achievable by brute force using dissipationless, N-body numerical simulations that treat the evolution of the cosmic density field under gravity only (e.g., Refs. [66; 67; 68; 69]). The most challenging obstacle to such precise predictions is the significant influence that baryonic processes have on the matter power spectra on scales of interest. This has been pointed out in both analytic [191; 199] and numerical studies [89; 155].

The problem posed by baryonic processes is severe because they cannot be modeled directly and all such calculations rely on relatively poorly-constrained, effective models that approximate the net, large-scale, effects of processes that occur on scales far below the numerical resolution that may be achieved with any simulation. Ref. [155] shows that the influences of baryons may be large compared with the statistical uncertainties expected of future surveys and that different treatments of baryonic physics lead to notably different matter power spectra on relevant scales. Ref. [197] studied a parameterized model for baryonic influences that could treat all of the simulations of Ref. [155] and showed that such a parameterized model could be calibrated self-consistently using weak lensing data alone, yielding meaningful constraints on both the dark energy and the effective models of baryonic physics. However, extending the results of Ref. [197] to consistency checks of general relativity requires care. Baryonic processes manifest as an inability to predict the evolution of the density field even for a fixed cosmic expansion history. This is, in part, what Ref. [197]

relied upon in their self-calibration program and this is precisely what proposed tests of the consistency of general relativity rely upon.

The strategy behind all efforts to study dark energy using gravitational lensing relies on the ability to produce reliable N-body simulations. In fact, most studies implicitly assume such a simulation campaign will be performed prior to any data analysis and proceed to estimate the power of forthcoming experiments using contemporary fitting formulae. In fact these fitting formulae are not sufficiently precise to treat forthcoming data, they may be subject to fundamental limitations [68; 69; 114; 154; 184], and, aside from a handful of specific cases, they have not been generalized to treat dark energy. A numerical campaign will be necessary to address dark energy, and addressing observables using dissipationless N-body simulations is a challenging, but tractable problem [67; 69]. The results of N-body simulation campaigns represent precise solutions to the idealized problem of computing the density or lensing fields in a variety of cosmologies absent baryonic physics. As with dark energy constraints, the strategy of self-calibrating the net influences of baryons is based on the assumption that such a set of simulations without baryons are available and that a simple set of prescribed corrections can be applied to the N-body results to describe baryonic effects. The reason is that direct numerical calculations that treat the physics of baryons are not achievable in the foreseeable future, given current computational limitations. This is the context of the present chapter.

In this chapter, we extend the results of Ref. [197] to address consistency checks of general relativity. We present a brief pedagogical discussion of the manner in which independent modifications to the distance scale of the universe and the growth rate of structure produce non-degenerate changes in weak lensing statistics, enabling weak lensing alone to serve as a powerful consistency check for general relativity. We then move on to discuss the interplay between modifications to convergence spectra caused by baryonic processes and the dark energy. We show explicitly that our current ignorance regarding the evolution of the baryonic component of the universe, its influence on the evolution of inhomogeneities of dark matter, and the process of galaxy formation places a severe limitation on our ability to test the consistency of general relativity. To be specific, we consider two tests. In one test, the growth of structure and the cosmic distance scale are assumed to arise from two different

effective equations of state $w_{\rm g}$ and $w_{\rm d}$, and we test our ability to rule out the hypothesis that the two are equal as they would be in general relativity. If unaccounted for, our limited ability to predict convergence spectra can lead to biases that drive these two parameters to disagree at levels as large as $\sim 8\sigma$ (depending upon details) when general relativity is the correct description of gravity. We perform a similar test on the gravitational growth index parameter γ, introduced in Ref. [111], and likewise find significant biases if baryonic effects are not treated. We show that it is possible to eliminate these biases by disregarding the small-scale shear information from forthcoming surveys, but the cost is a factor of $\sim 2-4$ degradation in the constraining power of such surveys. Lastly, we study the ability of forthcoming surveys to test general relativity and constrain baryonic processes simultaneously within a single data set. We show that this is a promising option as the biases can be eliminated at a cost of only $\sim 20\%$ in parameter constraints. This implies that significant effort should be devoted to developing robust parameterizations of the nonlinear evolution of structure as we begin to realize the next generation of imaging surveys. These methods should extend current techniques by including realistic descriptions of the effects of baryons that are not so general as to be completely arbitrary but do allow sufficient freedom so as to reflect our ignorance of the influence of the baryonic sector.

In the following section, we describe our methods, including our parameterizations of cosmological expansion and structure growth. In § 2.3, we present the results of our study. We begin in § 2.3.1 by illustrating the power of these consistency checks. This section contains a compilation of facts and figures found dispersed throughout the existing literature. In § 2.3.2, we show how neglect of baryonic processes may lead to biases that would mimic an inconsistency in general relativity in simpler treatments of weak lensing observables. In § 2.3.3, we present our results that show that additional parameter freedom may be added to predictions of the growth of structure to account for the influence of baryons and make such consistency checks robust to nonlinear processes. In § 2.4, we discuss our results and we summarize our efforts in § 2.5.

2.2 METHODS

2.2.1 Weak Lensing Observables

We treat weak lensing observations as a consistency check of general relativity. Aside from our parameterization of dark energy, we perform all calculations as described in detail in Ref. [197]. We summarize those calculations here and refer the reader to this reference for more detail. We consider number density-weighted convergence power spectra and cross-spectra from $N_{\rm TOM} = 5$ tomographic redshift bins as our primary observables:

$$P_\kappa^{\rm ij}(\ell) = \int {\rm d}z \frac{W_{\rm i}(z) W_{\rm j}(z)}{H(z) D_{\rm A}^2(z)} P(k = \ell/D_{\rm A}, z). \tag{2.1}$$

The five photometric redshift bins are spaced equally in redshift between a minimum photometric redshift of $z_{\rm p}^{\rm min} = 0$ and a maximum $z_{\rm p}^{\rm max} = 3$. Previous work showed that parameter constraints are saturated with $N_{\rm TOM} = 5$ and that further binning is unnecessary [119; 197]. We have verified that this remains so in the models we consider. In Eq. (2.1), $D_{\rm A}$ is the angular diameter distance, $H(z)$ is the Hubble expansion rate, $P(k, z)$ is the matter power spectrum at wavenumber k and redshift z, and $W_{\rm i}(z)$ are the lensing weight functions. Given the true redshift distribution of sources in bin i, ${\rm d}n_i/{\rm d}z$, the lensing weight is

$$W_{\rm i} = \frac{3\Omega_{\rm m} H_0^2}{2} (1+z) D_{\rm A} \int_z \frac{D_{\rm A}(z, z')}{D_{\rm A}(z')} \frac{{\rm d}n_i}{{\rm d}z'} \, {\rm d}z'. \tag{2.2}$$

We assume that the true redshift distribution of sources is

$${\rm d}n(z)/{\rm d}z = 4\bar{n} (z/z_0)^2 \exp[-(z/z_0)^2] / \sqrt{2\pi z_0^2}, \tag{2.3}$$

where $\bar{n}$ represents the total density of source galaxies per unit solid angle, which varies from survey to survey, and we assume $z_0 \simeq 0.92$ so that each survey has a median redshift of $z_{\rm med} = 1$. We assume that the probability of a photometric redshift $z_{\rm p}$ given a spectroscopic redshift z is Gaussian with a mean value of $z_{\rm p}$ given by z (no bias) and a dispersion $\sigma_z = 0.05(1+z)$, and compute ${\rm d}n_i/{\rm d}z$ for each photometric redshift bin as described in Ref. [119].

We estimate the constraining power of forthcoming weak lensing surveys using the Fisher information matrix. Indexing the observables of Eq. (2.1) by a single label, we write $\mathcal{O}_{\rm A} =$

$P_\kappa^{\rm ij}$, where each i, j map onto a unique A, and the Fisher matrix of weak lensing observables is

$$F_{\alpha\beta} = \sum_{\ell_{\rm min}}^{\ell_{\rm max}} (2\ell+1) f_{\rm sky} \sum_{\rm A,B} \frac{\partial \mathcal{O}_{\rm A}}{\partial p_\alpha} [C^{-1}]_{\rm AB} \frac{\partial \mathcal{O}_{\rm B}}{\partial p_\beta} + F^{\rm P}_{\alpha\beta}. \tag{2.4}$$

The p_α represent the parameters of the model, Greek indices label model parameters, capital Latin indices label unique observables, and lower-case Latin indices label tomographic redshift bins. $f_{\rm sky}$ is the fraction of sky covered by an experiment, $C_{\rm AB}$ is the covariance matrix of observables, and the sum begins at $\ell_{\rm min} = 2/\sqrt{f_{\rm sky}}$ and runs to some $\ell_{\rm max}$. We generally take $\ell_{\rm max} = 3000$ to ensure that we only use scales where assumptions of weak lensing and Gaussian statistics are valid [32; 44; 165; 185; 192]. However, we present many of our primary results (Figures 4, 5, 6, 7) as a function of $\ell_{\rm max}$ so results for alternative choices are simple to extract. Observed spectra contain both signal and noise, $\bar{P}_\kappa^{\rm ij} = P_\kappa^{\rm ij} + n_{\rm i}\delta_{\rm ij}\langle\gamma^2\rangle$, where $n_{\rm i}$ is the surface density of source galaxies in bin i and $\langle\gamma^2\rangle$ is the intrinsic source galaxy shape noise. We follow convention in setting $\langle\gamma^2\rangle = 0.2$ and allowing differences in shape noise between different observations to be absorbed into $n_{\rm i}$. The covariance between observables $\bar{P}_\kappa^{\rm ij}$ and $\bar{P}_\kappa^{\rm kl}$ is $C_{\rm AB} = \bar{P}_\kappa^{\rm ik}\bar{P}_\kappa^{\rm jl} + \bar{P}_\kappa^{\rm il}\bar{P}_\kappa^{\rm jk}$, where i and j map to A, and k and l map to B.

The inverse of the Fisher matrix is an estimate of the parameter covariance near the maximum of the likelihood. The measurement error on parameter α marginalized over all other parameters is

$$\sigma(p_\alpha) = [F^{-1}]_{\alpha\alpha}. \tag{2.5}$$

The second term in Eq. (2.4) incorporates Gaussian priors on model parameters. We assume modest priors on several cosmological parameters individually, so that $F^{\rm P}_{\alpha\beta} = \delta_{\alpha\beta}/\sigma_\alpha^2$, where σ_α is the assumed prior uncertainty on parameter α. Unless otherwise stated, when we refer to the uncertainty in a parameter or subset of parameters, we are referring to the uncertainty in the parameters under discussion after marginalizing over the remaining parameters of the model. This formalism also provides an approximation for biases in cosmological parameter estimators due to unknown, untreated, systematic offsets in observables. Let $\Delta\mathcal{O}_{\rm A}$ be the difference between the true observable and the prediction for that observable absent the systematic. The induced bias in the estimator of parameter α due to the neglect of the

systematic offset is

$$\delta p_\alpha = \sum_\beta [F^{-1}]_{\alpha\beta} \sum_\ell (2\ell+1) f_{\rm sky} \sum_{\rm A,B} \Delta\mathcal{O}_{\rm A} [C^{-1}]_{\rm AB} \frac{\partial \mathcal{O}_{\rm B}}{\partial p_\beta}. \tag{2.6}$$

We model the matter power spectrum using the phenomenological halo model [117; 157; 164] (for a review see Ref. [33]). We use the particular implementation of Ref. [197]. This model utilizes standard fitting formulae for halo abundance and halo bias [169]. It is known that existing fitting formulae are not yet sufficiently precise to treat forthcoming data sets [114; 184], but it is likely that this can be overcome. Our approach is premised on the idea that dissipationless simulation programs will be carried out to calibrate these quantities to the necessary precision (or some equivalent strategy that utilized the simulation data directly), and that baryonic processes are the only processes that are not treated with sufficient precision. The non-standard modification to the halo model that we consider concerns the distribution of matter within dark matter halos. Typically, it is assumed that on average the mass density within a halo is described by the standard Navarro, Frenk, & White (NFW) density profile [133], $\rho \propto (cr/R_{\rm 200m})^{-1}(1+cr/R_{\rm 200m})^{-2}$, where $R_{\rm 200m}$ is the halo virial radius which we define to contain a mean density of 200 times the mean density of the universe, the normalization is set by the fact that the mass profile must integrate to contain the total virial mass m within $R_{\rm 200m}$, and c is the halo concentration which sets the radius of the transition between the two power laws as $r_{\rm s} \sim R_{\rm 200m}/c$.

The standard practice is to set the average concentration of a halo of mass m according to a phenomenological law derived from dissipationless N-body simulations, such as

$$c(m,z) = c_0 [m/m_{\star,0}]^{-\alpha} (1+z)^{-\beta}, \tag{2.7}$$

where $c_0 \approx 10$, $\alpha \approx 0.1$, $\beta \approx 1$ [21; 120; 134], and $m_{\star,0}$ is the mass of a typical object collapsing at $z = 0$. We neglect the spread in halo concentrations at fixed mass because this is gives rise to a negligible effect on the scales we consider, and on smaller scales this dispersion is degenerate with an overall shift in the concentration-mass relation [32; 45]. The shortcoming of Eq. (2.7) is that it describes halos in dissipationless N-body simulations that neglect the physics of baryons. Fortunately, Ref. [197] demonstrated that adopting a modified concentration-mass law within the halo model suffices to model the convergence

power spectra predicted by the baryonic simulations of Ref. [155] to a level where biases in inferred dark energy parameters are less than 10% of their statistical uncertainties for forthcoming surveys. We use this as a starting point for our preliminary estimate of the influence of baryonic processes on tests of the consistency of general relativity through weak gravitational lensing.

2.2.2 Parameterized Tests of the Consistency of General Relativity

As we mentioned in the introductory section, the ideal scenario would be to test parameterized families of theories for modified gravity that make specific and unique predictions; however, this is difficult at present. Few if any viable alternatives to general relativity that contain a contemporary epoch of accelerated expansion have been identified. Moreover, weak lensing requires some treatment of the evolution of density perturbations beyond the linear order of perturbation theory, and the nonlinear evolution of perturbations in models of modified gravity are not completely specified, nor have they been studied thoroughly. We consider two different parameterizations that have been proposed to test for the consistency of general relativity within forthcoming data. These do not represent complete, entirely self-consistent descriptions of any particular phenomenological alternative general relativity. We consider them as a pragmatic step toward testing the null hypothesis of gravity described by general relativity in the case of modest departures from the standard gravity. In both cases, we explore deviations to the linear growth of perturbations but assume that the relation between linear and nonlinear perturbations is unchanged. In practice, this means that we assume that halos of dark matter form with the abundances and other gross properties that they otherwise would have in a standard-gravity treatment and continue to employ the halo model to predict convergence spectra [Eq. (2.1)] on nonlinear scales ($\ell \gtrsim 300$). Modified gravity models require some environmental dependence to the gravitational force law so that gravity may deviate from general relativity on large scales (low density) yet satisfy local constraints on deviations from general relativity (in the high density environment of our galaxy) [52; 74; 97; 98; 131; 132; 140]. We do not consider any such modifications because there is no comprehensive treatment of the nonlinear evolution of perturbations in such theories and

no phenomenological model akin to the halo model in the case of general relativity, though preliminary work in this direction has begun [140].

Our first test is to split the dark energy equation of state parameter into two distinct parameters. The first parameter, $w_{\rm g}$, is used in all calculations of the growth of density perturbations while the second, $w_{\rm d}$, is used in all calculations of the relationship between redshift and distance. The deflection of light is given by a sum of the Newtonian and curvature potentials and in the context of general relativity with zero anisotropic stress perturbation, these potentials are equal[1]. This equality means that perturbation growth and the distance-redshift relation are both determined by the evolution of the cosmic expansion rate $H(z)$, so that $w_{\rm g}$ and $w_{\rm d}$ should be equal in general relativity and that an indication to the contrary may be a sign of modified gravity. Of course, any observational or theoretical systematic errors that drive the inferred values of $w_{\rm g} \neq w_{\rm d}$ must be controlled and accounted for in order for this program to work. An alternative way to state our present aim is that we seek to assess the importance of the inability to predict the influence of the baryonic component of the universe on lensing power spectra for such consistency checks and explore a method to ensure the robustness of such a consistency check. In this first test, we consider our parameter set for dark energy to consist of the present dark energy density in units of the critical density $\Omega_{\rm DE}$, as well as the two equation of state parameters $w_{\rm d}$ and $w_{\rm g}$. We take no priors on these parameters and set their fiducial values to $\Omega_{\rm DE} = 0.76$, $w_{\rm d} = w_{\rm g} = -1$.

In the second case, we explore the gravitational growth index parameter γ, introduced by Linder in Ref. [111] and explored in Refs. [79; 112]. Linder showed that a natural separation between the expansion history of the universe and the growth of perturbations could be achieved by taking the evolution of an overdensity δ to be given by

$$\frac{{\rm d}\ln\delta}{{\rm d}\ln a} = \Omega_{\rm m}(a)^{\gamma}. \tag{2.8}$$

Our notation is such that $a = 1/(1+z)$ is the cosmic scale factor, $\Omega_{\rm m}$ with no argument of scale factor or redshift is the current density of non-relativistic matter in units of the critical

[1]We are admittedly a bit cavalier here as there are several different notational conventions in use, see Refs. [9; 74; 75; 115; 116; 129]; however, our contribution is not to explore the detailed behavior of the metric potentials in any specific theory (largely for lack of such options), so we forego a detailed discussion. In fact, we have already assumed the equality of the Newtonian and curvature potentials in order to derive Eq. (2.1) for the convergence power spectrum in terms of the power spectrum of density fluctuations and which would otherwise be written in terms of the power spectra of the metric potentials directly.

density, $\Omega_{\rm m}(a) = \Omega_{\rm m} a^{-3} H_0^2/H(a)^2$ with an explicit argument represents the evolution of the ratio of the matter density to the critical density, and $H_0 \equiv H(z=0) = 100h$ km/s/Mpc is the present Hubble expansion rate. Eq. (2.8) holds for perturbations independent of scale or cosmic expansion history (aside from the dilution of $\Omega_{\rm m}(a)$. In an enormous variety of dark energy models embedded within general relativity the growth index obeys [79; 111; 112]

$$\begin{aligned} \gamma &= 0.55 + 0.02[1 + w(z=1)], \qquad \text{for} \quad w < -1 \\ &= 0.55 + 0.05[1 + w(z=1)], \qquad \text{for} \quad w > -1, \end{aligned} \tag{2.9}$$

where $w(z)$ is the dark energy equation of state evaluated at redshift z. Deviations from this relation therefore, may indicate non-standard gravity. This treatment is accurate to better than a percent in specific cases such as the self-accelerating braneworld gravity model of Ref. [47] (so-called DGP gravity, which is now ruled out as an explanation of the cosmic acceleration, see Ref. [51], but still serves as a useful example of the utility of this parameterization), which yields $\gamma \simeq 0.69$, as well as scalar-tensor and $f(R)$ theories of gravity [111; 112]. In this case, we consider the expansion history to be dictated by dark energy with abundance $\Omega_{\rm DE}$ and a time-dependent effective equation of state $w(a) = w_0 + w_{\rm a}(1-a)$ [27], while perturbation growth is dictated by γ. We take no prior constraints on these parameters and perturb about the fiducial values $\Omega_{\rm DE} = 0.76$ (as in the first case), $w_0 = -1$, $w_{\rm a} = 0$, and $\gamma = 0.55$.

2.2.3 Cosmological Parameters and Future Lensing Surveys

In the previous section, we described the parameters we use to describe deviations from general relativistic gravity. To summarize, we take two test of the consistency of general relativity. In both cases, we assume dark energy with a present density of $\Omega_{\rm DE} = 0.76$ in our fiducial model. In the first case, we split the dark energy equation of state parameter into two pieces, one governing the growth of density inhomogeneities $w_{\rm g}$, and one governing the distance redshift relation $w_{\rm d}$. We take as a fiducial model $w_{\rm g} = w_{\rm d} = -1$ and assume no priors on either parameter. In the second case, we assume dark energy with an effective equation of state $w(a) = w_0 + w_{\rm a}(1-a)$, with fiducial values $w_0 = -1$ and $w_{\rm a} = 0$. In this

case we assume inhomogeneity growth parameterized by γ according to Eq. (2.8), and take the fiducial value of $\gamma = 0.55$ as in general relativistic gravity with a cosmological constant causing contemporary acceleration.

Beyond the parameters that describe the dark energy/modified gravity sector, we consider four other cosmological parameters that influence lensing power spectra and may be degenerate with dark energy parameters. These parameters and their fiducial values are: the non-relativistic matter density $\omega_{\rm m} \equiv \Omega_{\rm m} h^2 = 0.13$, the baryon density $\omega_{\rm b} = \Omega_{\rm b} h^2 = 0.0223$, the amplitude of primordial curvature fluctuations $\Delta^2_{\mathcal{R}} = 2.1 \times 10^{-9}$ (in practice we actually vary $\ln \Delta^2_{\mathcal{R}}$) at the pivot scale of $k_p = 0.05$ Mpc^{-1}, and the power-law index of the spectrum of primordial density perturbations $n_{\rm s} = 0.96$. We adopt conservative priors on each of these additional parameters that are comparable to contemporary constraints on each of these parameters individually [101]. To be specific, we take prior constraints of $\sigma(\omega_{\rm m}) = 0.007$, $\sigma(\omega_{\rm b}) = 10^{-3}$, $\sigma(\ln \Delta^2_{\mathcal{R}}) = 0.1$, and $\sigma(n_{\rm s}) = 0.04$. Lastly, we allow three parameters that describe the effective concentrations of halos to be determined internally from the same data. These are the parameters $c_0 = 13$, $\alpha = 0.05$, and $\beta = 1$ of the power-law halo mass-concentration relation in Eq. (2.7). The fiducial values we choose, $c_0 \approx 10$ and $\alpha \approx 0.1$, differ from those found in N-body simulations of structure formation [discussion after Eq. (2.7)]. We choose these parameters to reflect the results of the hydrodynamic simulations of Ref. [155].

We study constraints that one would expect from three forthcoming galaxy imaging surveys. As mentioned in § 2.2.1, we follow current convention by setting the intrinsic shape noise of source galaxies to be $\langle \gamma^2 \rangle = 0.2$ for all experiments and subsuming differences in the shape noise between surveys into differences in the effective number density of galaxies on the sky for each survey. We assume a redshift distribution of source galaxies as given in Eq. (2.3).

The most near-term survey that we explore is the dark energy survey (DES), which may begin operations in 2009 and have results as soon as 2011-2012[2]. For the DES, we take $f_{\rm sky} = 0.12$ and $\bar{n} = 15/\text{arcmin}^2$. We also consider an imaging survey that might be conducted as part of a future space-based mission such as the proposed Supernova Acceleration Probe

[2] http://www.darkenergysurvey.org

(SNAP)[3], which is a canonical example of a National Aeronautics and Space Administration, Beyond Einstein, Joint Dark Energy Mission probe[4], though the specifics of the mission have yet to be decided and other competitors include the Advanced Dark Energy Physics Telescope (ADEPT)[5] and the Dark Energy Space Telescope (DESTINY)[6]. For a SNAP-like probe, we take $f_{\rm sky} = 0.025$ and $\bar{n} = 100/\text{arcmin}^2$. Lastly, we consider a future ground-based imaging survey as might be carried out by the Large Synoptic Survey Telescope (LSST)[7]. We adopt $f_{\rm sky} = 0.5$ and $\bar{n} = 50/\text{arcmin}^2$ to describe the LSST survey. In all cases we consider only the statistical limitations of the surveys. The only systematic uncertainty we consider is the theoretical uncertainty associated with the inability to make precise predictions of the net influence of baryons on the lensing power spectra.

2.3 RESULTS

2.3.1 Consistency Checks

We begin by illustrating the power of weak lensing surveys to test the consistency of general relativity with forthcoming data. As we have mentioned above, the consistency checks work by comparing simultaneous constraints on cosmological distances and structure growth and determining whether they are consistent with a single, underlying, general relativistic model. Before moving on to parameter constraints we briefly discuss each of these effects individually. We will use the split dark energy parameterization model (with parameters $w_{\rm g}$ and $w_{\rm d}$) to illustrate the influence of modified distances and modified growth. Those readers with significant experience with dark energy phenomenology and proposed consistency checks of general relativity may like to proceed to the next section.

The influence of dark energy on the cosmic distance scale is demonstrated in Figure 1. The left panel of this figure shows the angular diameter distance as a function of redshift in

[3] http://snap.lbl.gov
[4] http://universe.nasa.gov/program/probes/jdem
[5] http://universe.nasa.gov/program/probes/adept
[6] http://www.noao.edu/noao/staff/lauer/destiny
[7] http://lsst.org

dark energy models with constant equations of state $w_{\rm d} = -1.1$, $w_{\rm d} = -1.0$, and $w_{\rm d} = -0.9$. The more negative the equation of state parameter, the more recent the prevalence of dark energy, and the more rapid is the current acceleration. Thus, for a fixed present Hubble expansion rate, lower values of $w_{\rm d}$ lead to larger angular diameter distances. The dependence of the angular diameter distance on the dark energy equation of state manifests in the lensing weight [Eq. (2.2)]; this dependence implies that the lensing weight varies with $w_{\rm d}$. The effects of this dependence are shown in the right panel of Fig. 1 for the particular case of the weight for the third tomographic bin containing $W_3(z)$, which contains sources with photometric redshifts in the range $1.2 \leq z_p < 1.8$. Notice that the lens weight itself (dashed line) extends to $z > 1.8$. In our model, sources in this bin have true redshifts that extend beyond $z = 1.8$ because of the relatively large dispersion in calibrated photometric redshifts. For more negative $w_{\rm d}$, the $W_3(z)$ is relatively increased and weighted relatively more toward low redshift. Heuristically, this can be understood by considering a hypothetical individual deflector. Relative to the deflector, the angle of deflection of a light ray is fixed by the deflector properties and the apparent change of position of the source as seen by an observer grows with angular diameter distance. The varying distances to fixed redshift also affect the mapping between multipole and wavenumber in the matter power spectrum at fixed redshift. Increased distances mean smaller wavenumber at fixed redshift, and as the power spectrum of density fluctuations is a declining function of wavenumber on scales of interest, this results in slightly stronger lensing.

Next, we demonstrate the dependence of the cosmic growth function on dark energy. Figure 2 shows the growth function $D(z)$ in our fiducial cosmology (dashed line), normalized such that $D(z) \to (1+z)^{-1}$ as $z \to \infty$. The dashed line in this figure shows the growth function in a flat cosmological model with no dark energy and $\Omega_{\rm m} = 1$. Comparing these two growth functions shows the dramatic suppression in the growth of density perturbations during the epochs where dark energy makes a significant contribution to the cosmic energy budget. In addition, we show the growth functions in cosmological models with $w_{\rm g} = -1.1$ and $w_{\rm g} = -0.9$ relative to the growth function in our fiducial model. The more positive the dark energy equation of state is, the earlier the acceleration begins and the earlier the growth of structure is quenched by the competing cosmic expansion. Conversely, more

negative values of the dark energy equation of state lead to more recent expansion and, as a result, more aggregate growth of density perturbations since the epoch of recombination. These features are all represented in Fig. 2.

The modifications to the growth of structure and cosmic distances manifest as changes in the observable power spectra. Figure 3 shows partial derivatives of three of the power spectra, P_κ^{22}, P_κ^{33}, and P_κ^{23}, as a function of parameters w_d, w_g, and $\ln\Delta^2_\mathcal{R}$. The derivatives with respect to the dark energy parameters are all negative. To reduce the range of the vertical axis, we actually display $-\mathrm{d}P_\kappa^{\mathrm{ij}}/\mathrm{d}\ln\Delta^2_\mathcal{R}$, so that it will also lie in the negative vertical half plane. Notice that variations in the observable power spectra due to changes in w_g and w_d have very different redshift and scale dependence. Consequently, we should not expect these parameters to exhibit a significant degeneracy and we would expect that shear power spectrum observations could constrain both parameters independently. Notice also that changes in w_g induce relatively larger shifts in the observable power spectra than do shifts in w_d. In the absence of additional information, this might be taken as evidence that weak lensing constraints on w_g should be more stringent than such constraints on w_d. In practice, w_g is strongly degenerate with the power spectrum normalization parameter $\ln\Delta^2_\mathcal{R}$. This can also be seen in Fig. 3. The dotted lines in Fig. 3 shows $-\mathrm{d}P_\kappa^{\mathrm{ij}}/\mathrm{d}\ln\Delta^2_\mathcal{R}$, and this quantity exhibits a similar redshift and scale dependence as the derivatives of the spectra with respect to w_g. Though not shown, w_g is also strongly degenerate with ω_b. The net result of these degeneracies is that w_d will be significantly more strongly constrained by weak lensing power spectrum observations than w_g. This discussion is closely related to the eigenmode analysis of Ref. [201] and serves as a qualitative demonstration that both w_g and w_d can be constrained by cosmic shear measurements, but the relative constraints on each parameter are sensitive to choices for external priors in the available additional parameter space.

Having reviewed the influence of independent changes in the cosmological distance scale and the rate of density fluctuation growth on observable weak lensing power spectra, we now turn to projections for the utility of general relativity consistency checks. We utilized the split dark energy equation of state parameterization for the previous illustrations, but we present results for both the w_d-w_g parameterization as well as the growth index (γ)

parameterization. Standard projections for the power of forthcoming consistency checks using weak lensing are summarized in Figure 4. Clearly such consistency checks can be quite powerful when definitive predictions can be made for the convergence power spectra. In the particular case of LSST, projected 1σ constraint on $w_{\rm d}$ is roughly $\sigma(w_{\rm d}) \simeq 0.02$, while the projected constraint on the gravitational growth index is $\sigma(\gamma) \simeq 0.04$, meaning that the LSST weak lensing program alone could rule out DGP gravity at the $\sim 3.5\sigma$ level. In the case of the $w_{\rm d}$-$w_{\rm g}$ parameterization, it is the equality of these two parameters that serves as a null hypothesis, so it is useful to look at the marginalized constraint on the combination $w_{\rm diff} \equiv w_{\rm d} - w_{\rm g}$. Transforming to this variable, we find a 1σ constraint $\sigma(w_{\rm diff}) = 0.04$ for LSST, $\sigma(w_{\rm diff}) = 0.11$ for SNAP, and $\sigma(w_{\rm diff}) = 0.14$ for DES.

Unfortunately, these standard projections assume perfect knowledge of nonlinear structure growth and ignore the influence of poorly-understood baryonic processes on the predictions of lensing spectra. Uncertainties due to baryonic processes and the process of galaxy formation are significant and will be the subject of the remainder of this section, beginning with the following section where we demonstrate that such effects can significantly bias estimators of dark energy parameters and diminish the utility of consistency checks if not accounted for.

2.3.2 Biases

The poorly-constrained baryonic processes are detrimental to any consistency check of general relativity because, if unaccounted for, they may induce biases in inferred cosmological parameters. In particular, if the estimators of dark energy parameters that we have introduced are significantly biased, it would be possible to conclude erroneously that data are inconsistent with general relativity. The relevant quantities to examine in order to assess the importance of baryonic effects are the biases in units of the statistical uncertainties of those parameters. If the biases are small compared to the statistical uncertainties, it will be unlikely to rule out the true model based on these biases, but as biases become comparable to or larger than statistical uncertainties it becomes increasingly likely that the true model may be ruled out based on the data.

In practice it is difficult to estimate what biases may be realized, because the true power spectra must be known in order to perform such a calculation (and, of course, if the true spectra were calculable there would be no bias). Fig. 5 gives an estimate of biases that may reasonably be realized for a variety of assumptions regarding the true convergence power spectra. We constructed these bias estimates as follows. Following the simulation analysis of Refs. [197], we assume that the primary influence of baryons can be accounted for in a halo model by adopting a non-standard halo concentration relation. It has been demonstrated that incorporating this additional freedom into the halo model enables accurate modeling of the results of non-radiative and dissipational hydrodynamic cosmological simulations and reduces expected biases in inferred dark energy parameters to acceptable levels [197]. The standard result for the concentration relation from gravity-only simulations is Eq. (2.7) with roughly $c_0 = 10$, $\alpha = 0.1$, and $\beta = 1.0$ [134]. We take this model to be our treatment of the results of N-body simulations. The halos in the simulations of Ref. [155] exhibit a $\sim 40\%$ enhancement in their concentrations in dissipational simulations relative to N-body simulations, so a choice of $c_0 \approx 14$ is an acceptable model for these simulation results. Unfortunately, our understanding of the evolution of baryons is poor and no simulation campaign can produce definitive results (it is this lack of definitiveness that is the primary problem). In particular, this boost is likely an over-estimate of any true shift due to galaxy formation because contemporary simulations exhibit baryonic cooling and star formation that is more efficient than allowed observationally [7; 95; 107; 142; 177]. As a consequence, baryonic processes must be treated in a way that accommodates a wide range of predictions. We illustrate reasonable levels of bias due to galaxy formation effects by computing the biases induced in dark energy parameter estimators when the convergence spectra are modeled with a fixed halo concentration-mass relation with $c_0 = 10$, but where the "true" value of this normalization ranges among $c_0 = 11, 12, 13, 14,$ and, 15. The lower end of this range is only a small shift relative to the standard concentration relation and is realized even in non-radiative hydrodynamic simulations where baryons cannot radiate their energy and form galaxies. The upper end of this rangeis slightly larger than the Ref. [155] simulation results and is near the maximum allowable by current lensing constraints at low redshift ($z \sim 0.2$) [121]. We note in passing that there are as yet weak observational indications of concentrations higher

than those predicted by dissipationless simulations [24; 188].

Fig. 5 shows biases in the inferred $w_{\rm g}$ and $w_{\rm d}$ induced by treating convergence power spectra that would be well described by an enhanced concentration mass relation, where c_0 can be any of $c_0 = 11, 12, 13, 14,$ or 15, with a concentration-mass relation fixed to have $c_0 = 10$. The objective is to provide reasonable estimates for biases that may be realized by not treating baryonic processes appropriately. If the Ref. [155] simulation results were a good representation of the "true" convergence power spectra, the biases would be near the top of the range in Fig. 5, roughly corresponding to the fourth most strongly biased model in each panel.

To see why it is useful to consider biases in dark energy parameter estimators as a function of scale may require some elaboration. The influence of baryons is greater on smaller scales, so a simple way to eliminate the uncertainty caused by baryons is to excise small-scale information from parameter estimation. A simplistic way to do this is to choose a maximum multipole $\ell_{\rm max}$ for the analysis and to disregard spectra beyond that maximum (Ref. [81] explores more sophisticated means to excise small-scale information, but they yield similar results in this context [197]). In Fig. 5, we show biases as a function of $\ell_{\rm max}$ in order to illustrate the reduced biases realized as a function of minimum scale (maximum-multipole).

Fig. 5 illustrates several points of interest. Most previous lensing analysis have assumed that one could utilize information to at least $\ell_{\rm max} \sim 10^3$ effectively. Fig. 5 shows that if one considers such small-scale information, estimators of dark energy equation of state parameters $w_{\rm g}$ and $w_{\rm d}$ will both be significantly biased relative to their true values. Notice also that they are generally biased in an *opposing* sense. This is of paramount importance for consistency checks where the goal is not necessarily to constrain either parameter individually, but to test the hypothesis that $w_{\rm g}$ and $w_{\rm d}$ are the same. For example, if $w_{\rm g}$ and $w_{\rm d}$ were always biased in the same way due to an unaccounted for systematic then it would not be possible to produce a reliable constraint on either parameter, but in this contrived example it would still be possible to test for their equality without accounting for the systematic. The biases in Fig. 5 are generally large compared to statistical uncertainties and suggest that consistency checks of general relativity can be compromised by uncertainty in galaxy formation physics and may lead to rejection of the hypothesis of general relativity even in models where this is

the true theory of gravity. The point where biases become acceptable ($\delta_w \ll \sigma_w$) varies from one experiment to the next. One could argue that for the LSST survey these biases are not acceptable for any $\ell_{\rm max} \gtrsim 300$.

As a rough criterion, Fig. 5 demonstrates that reducing the biases induced by the shifts in halo structure to acceptable levels requires taking a maximum multipole of no more than a few hundred. Of course, this reduction in information comes at a cost. The monotonically decreasing lines in Fig. 5 show the marginalized 1σ statistical uncertainties in $w_{\rm d}$ and $w_{\rm g}$ as a function of the maximum multipole $\ell_{\rm max}$ in units of the uncertainty when all information to an $\ell_{\rm max} = 3000$ is used. Notice that these uncertainties should be read relative to the right vertical axes in each panel. Overlaying the dependence of parameter uncertainties on this plot serves to demonstrate the cost of excising small-scale information. For LSST, the cost is large, a factor of $\sim 3-4$ in the uncertainties of both $w_{\rm d}$ and $w_{\rm g}$, because of the exquisite precision with which LSST could measure convergence spectra. This greatly reduces the effectiveness of an LSST-like data set to test the consistency of general relativity. DES is least affected by the uncertainty in galaxy formation because it surveys a comparably low number density of galaxies and makes the least use of small-scale information of any of the three experiments. SNAP is intermediate between the two with constraints degraded by roughly a factor of ~ 2.5.

Figure 6 shows the biased estimator for the gravitational growth index introduced in Ref. [111]. Our methods for computing biases and the presentation are analogous to that of Fig. 5. Of course, in this case the only parameter of importance is the gravitational growth index γ. Qualitatively, Fig. 6 shows results that are similar to that of Fig. 5. In particular, biases can be considerable compared to statistical uncertainties. Moreover, excising small-scale information comes at great cost in the statistical uncertainties of the dark energy parameters, with relative degradation of $\sim 2-4$ depending upon the choice of $\ell_{\rm max}$ and experimental parameters.

2.3.3 Calibration

In the previous section, we demonstrated the significance of the biases that would be induced in tests of general relativity if baryonic processes were to be ignored. We also showed the significant cost of excising information from the small scales where baryonic effects are most important in terms of degraded parameter uncertainties and thus degraded tests of the theory of gravity. What remains is to study whether such tests can still be performed utilizing high-multipole information if baryonic effects are treated in a parameterized way and calibrated simultaneously with the cosmological parameters. Though there is still significant work to be done, preliminary indications are that baryonic effects can indeed be treated by modifying the predictions of dissipationless N-body simulations [155; 155]. In fact, the necessary modification is to treat the relation between halo mass and halo concentration as free, and we have already utilized such modifications to estimate parameter biases in the previous section.

The next step is to determine the degree to which treating both halo structure and cosmology as uncertain, in order to eliminate the biases of the previous section, and fitting both to forthcoming data will degrade our ability to test general relativity. If the degradation is significantly less than that incurred by simply excising small-scale information, then an appropriate strategy to adopt in the run-up to the next generation of imaging surveys would be to spend significant effort understanding the phenomenology of structure growth in the nonlinear regime in a variety of different effective models for the evolution of the baryonic sector. It might then be possible to "self-calibrate" the influence of baryons, simultaneously yielding insight into the law of gravity on large scales and the small-scale physics of galaxy formation. Alternatively, if the degradation is comparable to the degradation incurred by excising small-scale information, such a program would likely be futile and it would be wiser to excise small-scale information until such time as definitive predictions for the influence of baryonic processes on lensing power spectra can be made.

There is already reason to be optimistic that some self-calibration can be achieved. Ref. [82] and Ref. [4] demonstrated the robustness of weak lensing to modest multiplicative and additive errors in shear measurements. In addition, Ref. [197] have already demon-

strated robustness of weak lensing constraints on dark energy to some uncertainty in the power spectrum at high wavenumbers. A large part of potential to carry out self-calibration programs within weak lensing data can be gleaned from the eigenmode analysis of Ref. [201]. The robustness of weak lensing constraints on dark energy is due in part to the fact that weak lensing provides exquisite constraints on distance measures that are relatively insensitive to dark energy and can be brought to bear on other systematics without degrading dark energy constraints.

In this section, we present the results of a self-calibration exercise where we treat cosmological parameters as before and introduce additional parameter freedom to describe the influence of baryons. To be specific, we treat the halo concentrations according to Eq. (2.7), but allow the parameters c_0, α, and β to be free. It is already possible to offer a guess at the promise of such a self-calibration exercise. Fig. 3 shows the derivatives of three of our convergence power spectra P_κ^{22}, P_κ^{23}, and P_κ^{33} with respect to the parameters $w_{\rm d}$, $w_{\rm g}$, and $\ln\Delta^2_{\mathcal{R}}$. Compare this to the shifts in P_κ^{33} depicted in the upper, left panels of Fig. 5. Modifying the concentration relation causes significant increases in power at small scales (high multipoles). As a result, there is a stronger scale dependence associated with changes in concentration parameters than there is associated with changes to $w_{\rm d}$, $w_{\rm g}$, γ, or $\Delta^2_{\mathcal{R}}$. This is encouraging because it suggests that if the scale dependence can be adequately modeled, the parameters describing baryonic processes can be extracted independently without significantly degrading consistency checks of general relativity.

We show the results of the self-calibration exercise in Figure 7. In this panel we show constraints in the case of self calibration of halo structure $\sigma^{\rm self-cal}(\ell_{\rm max})$, in units of the parameter uncertainties that would be quoted in the standard case of perfect knowledge of halo structure, $\sigma^{\rm standard}(\ell_{\rm max})$ (shown in Fig. 4, Fig. 5, and Fig. 6). In particular, we show how the parameter degradation varies with $\ell_{\rm max}$ because it is both instructive and it illustrates the balance between excising data and self-calibrating. Figure 7 shows results for both the $w_{\rm d}$-$w_{\rm g}$ parameterization as well as the gravitational growth index (γ) parameterization.

It may seem odd that the functions in Fig. 7 do not decrease monotonically with increasing $\ell_{\rm max}$. These relations are not monotonic functions of $\ell_{\rm max}$ because they represent the ratio of the error realized in the self-calibration calculation compared to the errors computed

in the limit of perfect knowledge of the influence of baryons. From Fig. 5 and Fig. 6, the constraints on all parameters are rapidly decreasing functions of $\ell_{\rm max}$. This is sensible; as more information is added constraints should improve and not degrade. Likewise, in the case of self-calibration the absolute constraints on each parameter decrease rapidly with increasing $\ell_{\rm max}$. What is not monotonic is the relative degradation of these constraints as a function of $\ell_{\rm max}$.

The scale dependence of the parameter degradation is not surprising. From the power spectra in the upper, left panel of Fig. 5 and the derivatives in Fig. 3 it is evident that each of the parameters of interest induces a scale-dependence on the observed spectra, so some scales are more effective at constraining dark energy parameters and calibrating halo structure than others. Though the details depend upon the experiment, and in particular the statistical weight an experiment places on a particular multipole, there are sensible general trends depicted in Fig. 7. In the limit of low-ℓ, halo structure is unimportant and results in the self-calibration calculation approach the standard results. This is least evident for LSST because LSST has sufficient $f_{\rm sky}$ to make precise power spectrum measurements even for $\ell \sim 300$. We do not extend these plots to $\ell < 300$ because at this value of ℓ parameter biases are already at or near acceptable levels and so there is no need to relegate ourselves to such low multipoles in any data analysis. With increasing multipole, the halo structure becomes ever more important and is mildly degenerate with dark energy. Further increasing $\ell_{\rm max}$ beyond $\sim$ several hundred introduces the strong scale dependence that halo structure imparts on the convergence spectra and this breaks a degeneracy between dark energy and halo structure. For this reason, the degradation decreases rapidly from $\ell_{\rm max} \sim 500$ to $\ell_{\rm max} \sim 2000$.

The most important aspect of Fig. 7 is that the degradations are relatively small if we utilize all of the information to $\ell_{\rm max} \gtrsim 10^3$. Degradations are scale dependent, but consider the level of degradation at $\ell_{\rm max} = 3000$ for definiteness. After calibrating halo structure, the $w_{\rm d}$ constraint is degraded by $\sim 34\%$ for LSST and roughly $\sim 18\%$ for DES and SNAP. The $w_{\rm g}$ constraint is degraded by $\sim 18\%$ for LSST and less than 10% for LSST and SNAP. The relevant combimation $w_{\rm diff} = w_{\rm d} - w_{\rm g}$ is degraded by a maximum of $\sim 36\%$ for LSST and only about $\sim 13\%$ for SNAP and DES. To appreciate the utility of self-calibration in this instance, we should compare these numbers to the information loss associated with excising

small scale information in Fig. 5. For LSST, excising high-ℓ information to eliminate biases requires taking $\ell_{\rm max} \sim 300$ with a corresponding $\sim 300\%$ increase in the constraints on $w_{\rm d}$ and $w_{\rm g}$. So long as baryonic effects can be modeled in this way, the more sensible strategy is clearly to use all available information and calibrate the physics of galaxy formation as encoded in effective halo concentrations. The rightmost panel of Fig. 7 shows a similar result for the gravitational growth index γ. In particular, σ_γ is degraded by only $\sim 20\%$ after self calibration as compared to $\sim 200\%$ as would be required by the excision of small scales (see Fig. 6).

For completeness, we give a revised set of constraint contours in the $w_{\rm d}$-$w_{\rm g}$ plane in Fig. 8. After calibration, the contours are slightly expanded, but their orientation in this plane is only slightly changed. By comparison, eliminating small-scale information corresponds to a dramatic loss of constraining power both because the area of the constraint contour in the $w_{\rm d}$-$w_{\rm g}$ plane grows and because its orientation changes such that the most degenerate combination becomes more nearly perpindicular to the line $w_{\rm d} = w_{\rm g}$. The overall conclusion remains that self calibration is the appropriate strategy to adopt.

2.4 DISCUSSION

We have extended recent studies of the influence of galaxy formation on weak lensing observables (in particular, Ref. [197]) to address cross-checks of the consistency of general relativity with weak lensing. Weak lensing is a key ingredient in such programs because weak lensing observables are sensitive to both the cosmological distance scale and the evolution of potential inhomogeneities, whereas the Type Ia supernovae distance-redshift relation and baryonic acoustic oscillation measurements are sensitive only to geometry (though it may be possible to exploit the lensing of superovae while performing the distance-redshift test, see Refs. [42; 125; 195] or to combine distance measures with cluster counts). Proposed methods to utilize weak lensing to constrain dark energy and to place limits on deviations from general relativity on cosmological scales assume that correlations on scales as small as a few arcminutes ($\ell \sim 10^3$) can be brought to bear on the cause of cosmic acceleration.

Such small-scale information is important to these programs. Relegating consideration to larger scales can significantly degrade the constraining power of forthcoming experiments (see Fig. 5 and Fig. 6) by nearly a factor of ~ 3 decrease in the effectiveness of tests for deviations from general relativity.

One of the anticipated challenges of using this small-scale information is that it is difficult to predict observables on these scales because the physics that governs the evolution of the baryonic component of the Universe is poorly understood and affects these observables in important ways [89; 155; 191; 199]. In Fig. 5 and Fig. 6, we show that using such data to test general relativity may introduce potentially large biases in the inferred values of cosmological parameters that can lead to ruling out general relativity when it is in fact the true theory of gravitation. In these figures, we give a range of estimates of potential biases that may be realized. These results are motivated by the specific results of the simulations of [155]. The biases can be many times the statistical uncertainties in these parameters, which is clearly a serious problem. We reiterate here that we give a range of biases that may be realized precisely because it is difficult to predict the precise influences of baryons. We have used simulation results as guidance and explored a range of concentration relations that are consistent with current observations (e.g., Ref. [121]). We have not included information from additional observables that can be brought to bear on this problem. For example, supernova Ia and baryon acoustic oscillation data can constrain the cosmic distance-redshift relation and be combined with lensing to produce stronger consistency checks. However, these additional constraints on cosmological distances would only make the discord between structure growth and distance measures more egregious and drive larger biases, making the problem more severe.

Faced with a large systematic error, a common practice is to degrade the experiment in some way so as to derive robust results from the data and reduce sensitivity to the systematic error. Baryonic influences are scale dependent, so one way to degrade the experiment and eliminate these biases is to disregard small-scale information by restricting the analysis to $\ell_{\rm max} \sim$ a few $\times\, 10^2$. However, in Fig. 5 we have shown that this comes at the cost of nearly a factor of ~ 3 decrease in the effectiveness of tests for deviations from general relativity (more sophisticated analyses, such as those in Ref. [81], yield similar results).

As an alternative to excising small scale information, simulations suggest that we can understand the form of the baryonic influence and correct for it using a parameterized model. To the degree that this form is not exactly degenerate with parameterizations of dark energy or deviations from general relativity, this enables the utilization of some of the information on small scales. We addressed these biases by including 3 baryonic parameters, α, β, and c_0, that model the effective concentration of dark matter halos as specified in Eq. 2.7. This choice of parametrization is motivated by Refs. [155; 197], wherein the authors demonstrated that modifying halo concentrations allows one to model account for the net influences of baryonic processes on scales relevant to weak lensing parameter estimation. In particular, models in which the N-body results are corrected by a modified halo concentration relation track the convergence spectra from galaxy formation simulations out to $\ell \sim 5000$ accurately enough to reduce the biases in inferred dark energy parameters to levels below the statistical uncertainty of forthcoming experiments. Within this framework, we studied the ability of forthcoming imaging surveys to self-calibrate the values of the baryonic parameters and simultaneously use this information to check the consistency of general relativity.

In Fig. 7 we have illustrated the effectiveness of our method by using the Fisher Matrix formalism to estimate the constraints on γ and $w_g - w_d$ that will be obtained from future photometric surveys. From this figure, it is evident that the constraints on the consistency of general relativity degrade by just $\sim 30\%$ relative to the limit of perfect knowledge of nonlinear evolution of the gravitational potential. In Fig. 8 the constraints on the consistency of general relativity obtained through our methods are much tighter than those from analyses excising small scale information. In light of our findings, we conclude that self-calibration of some of the specific influences of baryonic physics may be an appealing alternative to disregarding the information contained on scales smaller than $\ell \sim 300$.

Our outlook for self-calibration of uncertain baryonic physics is quite positive; however, there are several important caveats to our study and a number of additional studies that must be undertaken in order to bring such a program to fruition. First, this correction for the physics of galaxy formation has only been applied to the simulations of Ref. [155]. As the physics of baryons is highly uncertain, a more comprehensive exploration of viable alternative models is needed in order to validate such a correction term. In fact, it is not reasonable to

suppose that the influences of baryons should be strictly confined to the internal structures of halos (see Ref. [175] for changes in the halo mass function) and other effects may prove important. Nevertheless, our study indicates that such an exploration is a fruitful pursuit.

Second, any such tests have important limitations. The self-calibration program that we have explored should guard against the possibility of ruling out general relativity when it is the correct theory of gravity. In our particular study, we have not explored complete, self-consistent phenomenologies for modified gravity. Rather, we have only explored proposals that are already in the literature in which additional parameters are introduced which have known, fixed values in general relativity and explored the ability of forthcoming surveys to constrain these parameters. In the event of a detection of a deviation from general relativity such tests provide limited information about the character of the correct, alternative theory. In large part, this is due to the fact that modified gravity models must have some scale dependence so that they can simultaneously drive accelerated cosmic expansion yet satisfy small-scale bounds on the theory of gravity [52; 74; 97; 98; 131; 132; 140]. If a modified gravity theory, including nonlinear evolution, were fully specified and explored and if galaxy formation processes were more completely understood, it would be possible to draw more specific conclusions from forthcoming imaging survey data.

Lastly, we note that we operate under the assumption that dissipationless N-body simulations will effectively calibrate the properties of the lensing field absent baryonic effects. In our halo model approach, this is tantamount to a calibration of dark matter halo abundance and halo clustering that is so accurate as to introduce negligible uncertainties. This is a common premise to all such studies of dark energy constraints as it is known that the density field is not yet calibrated with sufficient accuracy to analyze forthcoming data [80; 154; 184]. Our outlook is that it is plausible that a brute force simulation campaign can achieve sufficient accuracy in this regard, though it is certainly possible that reaching this goal may be complicated by several factors that have not yet been thoroughly studied. For example, the phenomenologies we explore here are relatively simple, as they treat modified gravity only as a modification to the growth function. Ultimately, it will be of interest to study more specific and more complete phenomenological models of modified gravity that may lead to large corrections to the density field on scales of several Mpc. Additionally, even in the absence of

corrections necessitated by modified gravity it is possible that a program of N-body simulations coupled with a calibration of baryonic effects may encounter unforeseen obstacles. In either case, the assumption that an adequate suite of N-body simulations will address the density field (in our parameterization halo bias and halo abundance) with sufficient precision so as to render errors in the predicted density field negligible in the dissipationless case may not be valid. In such cases, it would be necessary to allow for adequate model freedom to account for this uncertainty. In the language of this chapter, this might amount to internal calibration of halo abundance and halo bias (presumably with significant prior information) as well. Such additional freedom would lead to further degradation in dark energy and/or modified gravity parameters. We have chosen to limit the scope of our work along these lines and not studied this more complicated case.

2.5 SUMMARY

We have explored the ability of forthcoming weak lensing surveys to perform a consistency check on general relativity as the correct theory of gravity in light of recent uncertainties in predicting lensing observables on scales smaller than $\sim$ 10 arcmin due to the poorly-understood evolution of the baryonic component of the Universe. After a pedagogical demonstration of the utility of such consistency checks in § 2.3.1, we present the following results.

1. Conducting a study of the consistency of general relativity by analyzing weak lensing observables from a forthcoming imaging survey (such as DES, JDEM/SNAP, or LSST) out to scales as small as $\ell \sim 10^3$ could potentially lead to inferred cosmological parameters that are biased by many times the statistical uncertainty of such surveys if baryonic effects on the nonlinear evolution of the gravitational potential are ignored. The problem is severe in that *they could lead to rejection of the null hypothesis of general relativistic gravity, even when it is the true theory.* Specific results depend upon the assumed survey properties and the importance of baryonic processes.
2. Disregarding small-scale information beyond $\ell \sim$ a few $\times 10^2$ significantly reduces these biases because of the strong scale-dependence of the influence of baryonic physics on

weak lensing observables. However, excising this information degrades the constraining power of forthcoming imaging surveys by a factor of ~ 3.

3. As an alternative to excising small-scale information, we explore a program of self-calibrating a phenomenological model of baryonic processes that is motivated by recent numerical simulations. With such a model small-scale information out to $\ell \sim 10^3$ can be exploited, and biases can be reduced to levels below those of the statistical uncertainty of forthcoming experiments. As shown in Fig. 7 and Fig. 8, this method provides substantially tighter constraints on the consistency of general relativity than analyses that disregard small-scale information.
4. The program to calibrate the baryonic correction internally allows small-scale information to be used to constrain gravity and guards against ruling out general relativity when it is the true theory. In the case of internal calibration, the limits provided by general relativity consistency checks are degraded by only $\sim 30\%$ relative to the limits that would be achieved if nonlinear structure formation were perfectly understood. Moreover, interesting limits on galaxy formation models are produced as a "byproduct."

These results should be carefully considered in preparation for the weak lensing surveys that will be conducted in the next decade. At the very least, it should be clear that issues regarding the baryonic influences on lensing observables will need to be dealt with, but that uncertainty in this regime may not necessarily have dramatic consequences on the ability of imaging surveys to illuminate the cause of cosmic acceleration. It seems likely that a program of internal calibration of baryonic physics will allow lensing surveys to achieve stringent limits on gravity and dark energy. In order to ensure that this is the case, a comprehensive theoretical program will need to be undertaken to better understand viable alternatives for the net influences of baryons on observable power spectra.

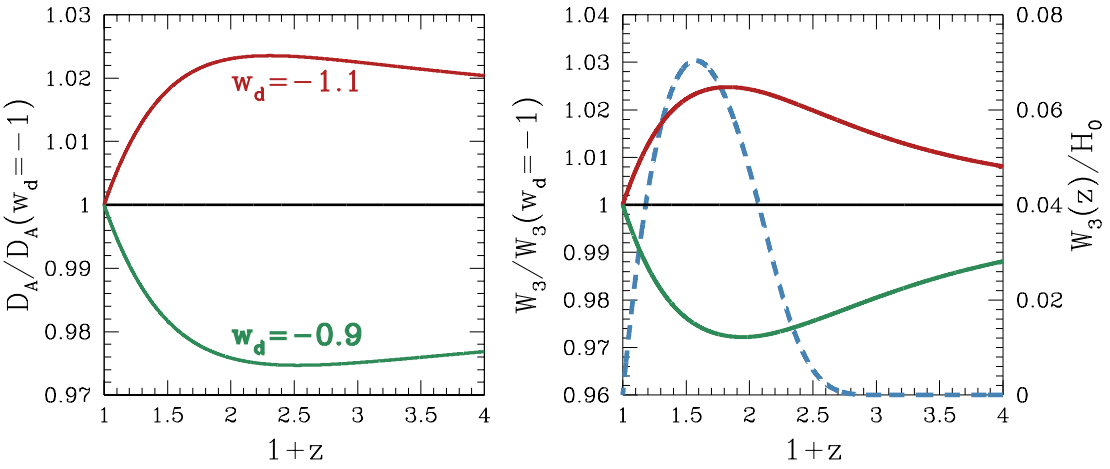

Figure 1: The influence of the dark energy equation of state parameter on the relationship between angular diameter distance and redshift. The *left* panel shows the angular diameter distance as a function of redshift for models with $w_{\rm d} \neq -1$, in units of the angular diameter distance in our fiducial model with $w_{\rm d} = -1$. The *upper* line shows $D_{\rm A}$ with $w_{\rm d} = -1.1$ and the *lower* line shows $D_{\rm A}$ with $w_{\rm d} = -0.9$. The *right* panel shows the net influence of this distance change on the lensing kernel of Eq. (2.2). The *dashed* line in this panel should be read against the right, vertical axis and represents the absolute weight $W_3(z)/H_0$. The solid lines show the particular example of the weight for the third tomographic bin $W_3(z)$ relative to its value in the $w_{\rm d} = -1$ case and should be read against the left vertical axis. The correspondence with $w_{\rm d}$ is as in the *left* panel.

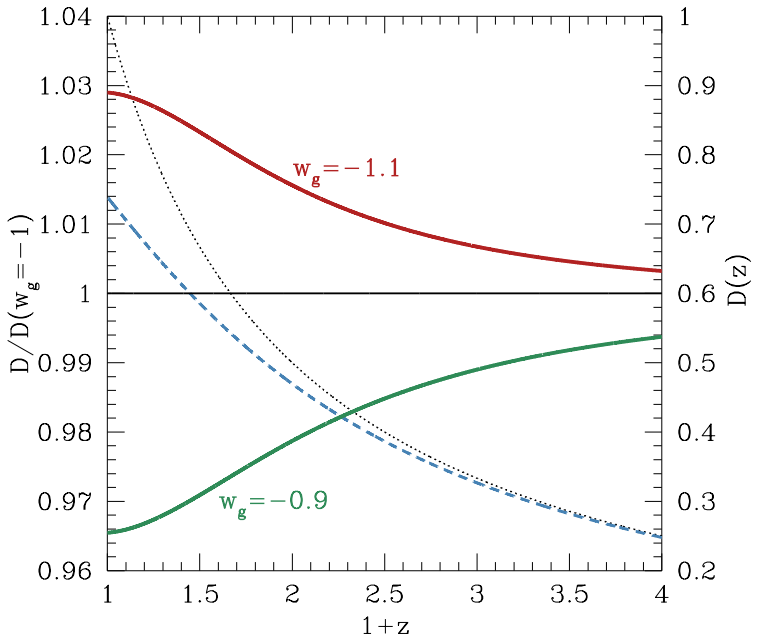

Figure 2: The influence of the dark energy equation of state parameter on the linear growth of structure over an observationally-relevant range of redshifts for forthcoming imaging surveys. We plot the linear growth function for dark energy with $w_{\rm g} = -1.1$ (*upper* line) and $w_{\rm g} = -0.9$ (*lower* line) as relative to the growth function in a cosmological constant model with $w_{\rm g} = -1.0$. These relative shifts should be read against the left, vertical axis. In analogy with Fig. 1, we plot the absolute growth function for the standard case of $w_{\rm g} = -1$ as the *dashed* line, which should be read against the right, vertical axis. The *dotted* line shows the growth function in a flat cosmological model with $\Omega_{\rm m} = 1$ and $\Omega_{\rm DE} = 0$ (and should again be read against the right, vertical axis). Comparing this growth rate to the growth rate in the fiducial model with $\Omega_{\rm DE} = 0.76$ illustrates the suppression of growth caused by the contemporary epoch of accelerated expansion.

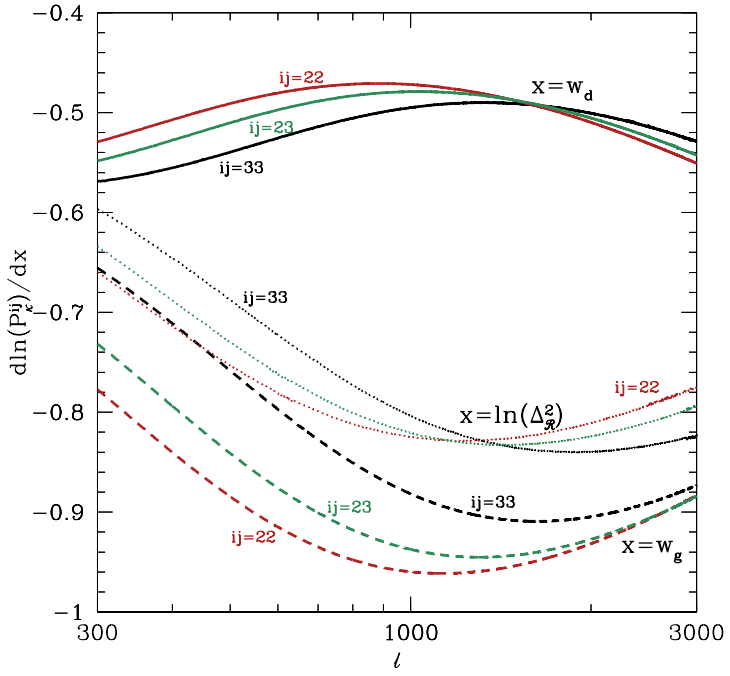

Figure 3: Partial derivatives of tomographic weak lensing power spectra with respect to cosmological parameters. We show derivatives of each of the power spectra $P_\kappa^{\rm ij}(\ell)$, with $ij = 22, 23, 33$ with respect to the three parameters $w_{\rm d}$, $w_{\rm g}$, and $\ln \Delta^2_{\mathcal{R}}$. The *solid* lines show derivatives with respect to $w_{\rm d}$. From bottom to top at left, the *solid* lines are $\partial \ln P_\kappa^{\rm ij}/\partial w_{\rm d}$ with $ij = 33, 23, 22$. The *dashed* lines represent derivative with respect to $w_{\rm g}$. From bottom to top at left, the *dashed* lines show $\partial \ln P_\kappa^{\rm ij}/\partial w_{\rm g}$ with $ij = 22, 23, 33$. The *dotted* lines are derivatives with respect to the power spectrum normalization parameter $\ln \Delta^2_{\mathcal{R}}$. From bottom to top at left, the *dotted* lines show $-\partial \ln P_\kappa^{\rm ij}/\partial \ln \Delta^2_{\mathcal{R}}$ with $ij = 22, 23, 33$. The additional negative sign for the derivatives with respect to $\ln \Delta^2_{\mathcal{R}}$ is designed to reduce the dynamic range on the vertical axis.

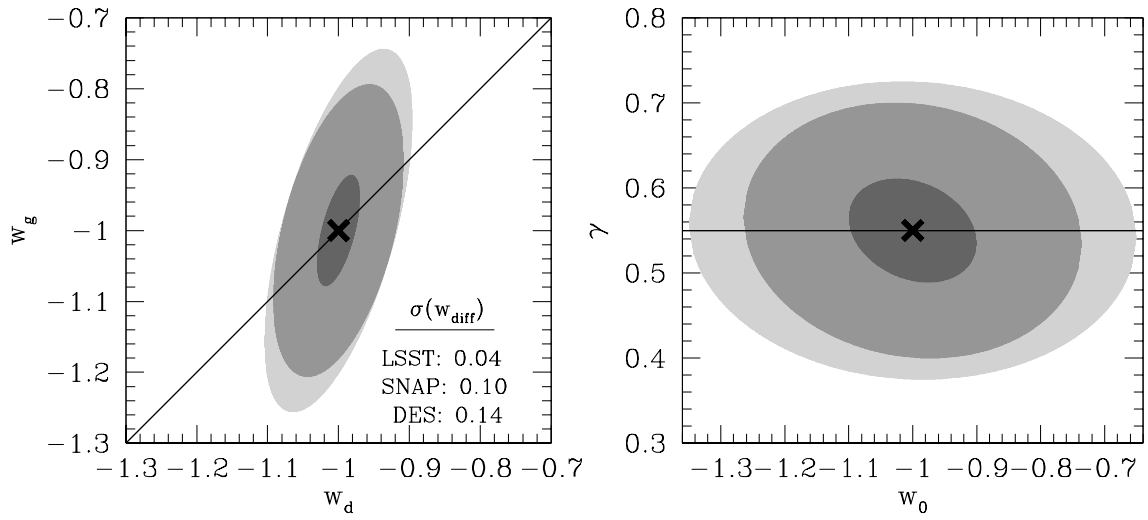

Figure 4: The capability of forthcoming surveys to perform consistency checks of general relativity. The *left* panel shows projected 1σ constraint contours in the $w_{\rm d}$-$w_{\rm g}$ plane for the LSST (innermost contour), DES (outermost contour), and SNAP weak lensing (middle contour) experiments. The diagonal line in this panel delineates $w_{\rm g} = w_{\rm d}$. The constraints on the parameter $w_{\rm diff} \equiv w_{\rm d} - w_{\rm g}$ from each experiment are given in the *lower, right* portion of the panel. The *right* panel shows constraints in the gravitational growth index parameterization in the w_0-γ plane. The three contours represent the LSST, DES, and SNAP weak lensing experiments as in the *left* panel.

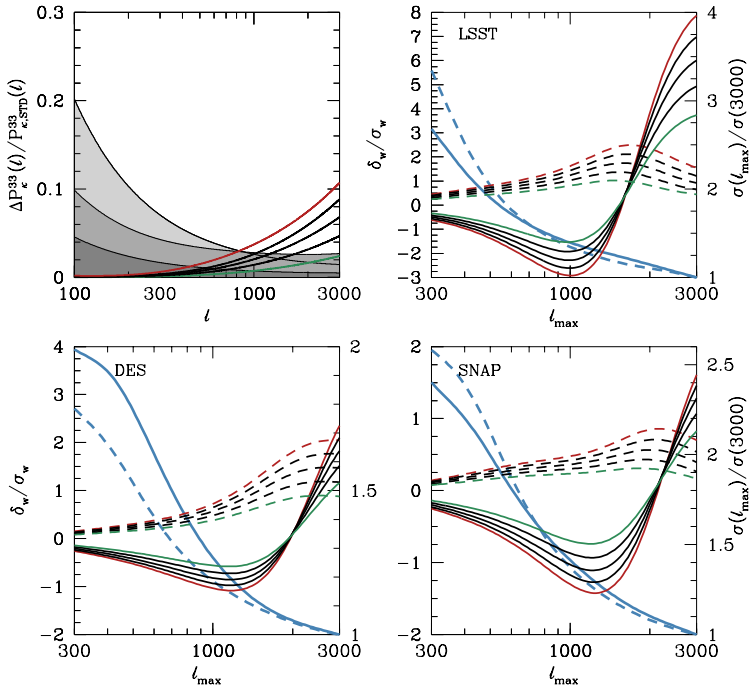

Figure 5: The *upper, left* panel shows the the effect of modified halo structure on convergence power spectra. Each of the lines that increases with multipole represents the relative change in a convergence power spectrum of sources in our third tomographic bin ($1.2 \leq z_p < 1.8$) $P_\kappa^{33}(\ell)$, in models with modified halo structure relative to that of the standard case. We represent the standard case using a halo model where halo concentrations are given by Eq. (2.7) with $c_0 = 10$, $\alpha = 0.1$, and $\beta = 1.0$. We represent models with modified halo structure by taking $c_0 = 11, 12, 13, 14,$ and, 15 from bottom to top. The shaded bands show the statistical errors on $P_\kappa^{33}(\ell)$ expected from forthcoming SNAP, DES, and LSST surveys from *top to bottom at left*. The other three panels show the biases in estimators $w_{\rm d}$ (*solid*) and $w_{\rm g}$ (*dashed*), in units of the statistical uncertainties in these parameters, as a function of the maximum multipole used in parameter estimation. These should be read against the *left* vertical axes. Each panel shows forecasts for a specific survey. The most biased cases correspond to $c_0 = 15$. The decreasing functions of $\ell_{\rm max}$ in each panel show the statistical uncertainty in $w_{\rm d}$ and $w_{\rm g}$ as a function of $\ell_{\rm max}$ relative to the error if all information to $\ell_{\rm max} = 3000$ were used. These lines should be read against the *right* vertical axes.

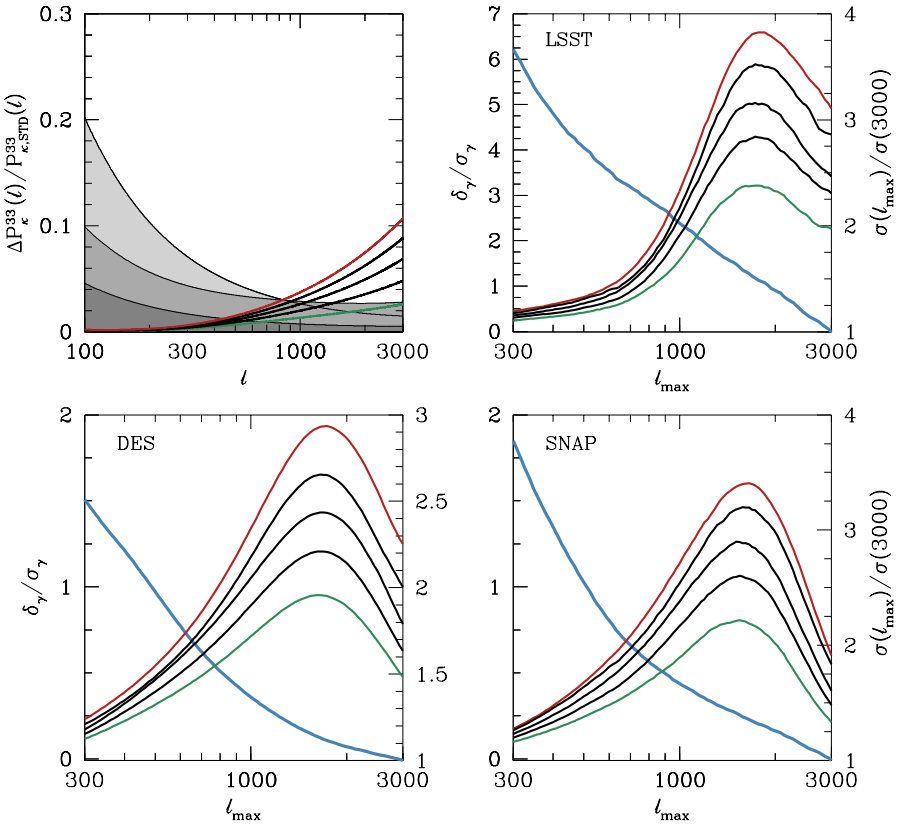

Figure 6: Estimates of biases in the estimator of the γ parameter of the gravitational growth index formalism proposed in Ref. [111]. The treatment of baryonic modifications to the convergence power spectrum is as in Fig. 5 and the *upper, left* panel is identical to the *upper, left* panel of Fig. 5. We repeat it here for convenience. The remaining panels show biases projected for the LSST, SNAP, and DES weak lensing surveys as indicated. Value of the biases should be read relative to the *left* vertical axes. The biases are given in units of the projected statistical uncertainty in γ and the most (least) biased case corresponds to a model with $c_0 = 15$ ($c_0 = 11$). The intermediate lines correspond to the intermediate values of c_0 with bias increasing with c_0. The monotonically decreasing lines in each panel represent the 1σ statistical uncertainties in γ as a function of maximum multipole $\ell_{\rm max}$. These are shown in units of the uncertainties in gamma achieved by considering all information to $\ell_{\rm max} = 3000$ in order to illustrate the relative degradation in constraints caused by excising small-scale information.

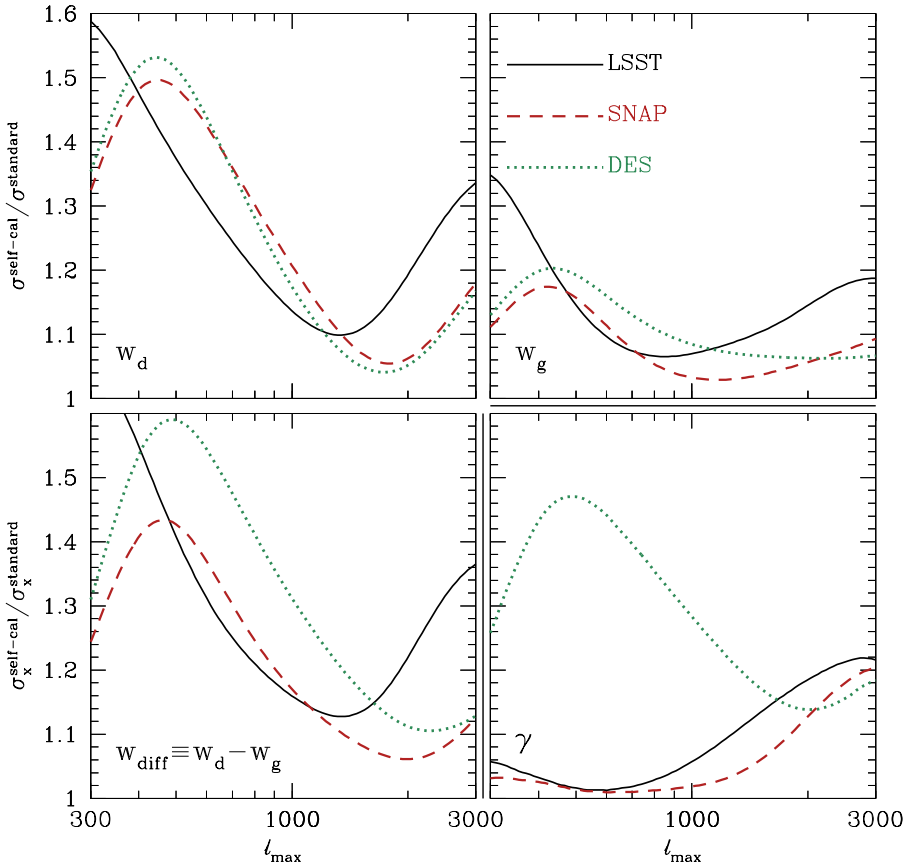

Figure 7: Results of self-calibrating halo structure simultaneously with tests for modified gravity on large scales. As with Fig. 5, we present results as a function of the maximum multipole considered in the parameter extraction analysis $\ell_{\rm max}$. The vertical axis is the error on the stated model parameter in the self-calibration case in units of $\sigma_X^{standard}$: the error on parameter X calculated by standard parameter forecasts that appear in the literature, where halo structure is assumed to be known perfectly. Each panel contains three lines for each set of experimental parameters we explore. The *solid* lines correspond to an LSST-like experiment, the *dashed* lines correspond to a SNAP-like experiment, and the *dotted* lines correspond to a DES-like experiment. Each panel focuses on a different parameter. The *upper, left* panel shows $w_{\rm d}$, the *upper, right* panel shows $w_{\rm g}$, the *bottom, left* panel shows the difference $w_{\rm diff} = w_{\rm d} - w_{\rm g}$, and the bottom, right panel shows the gravitational growth index γ. Note that $w_{\rm d}$ and $w_{\rm g}$ are the parameters of our "dark energy split" model and are constrained simultaneously, and that the relevant constraint for tests of gravity is the constraint on $w_{\rm diff}$. The γ parameter is from an entirely distinct parameterization in which dark energy is independently marginalized over.

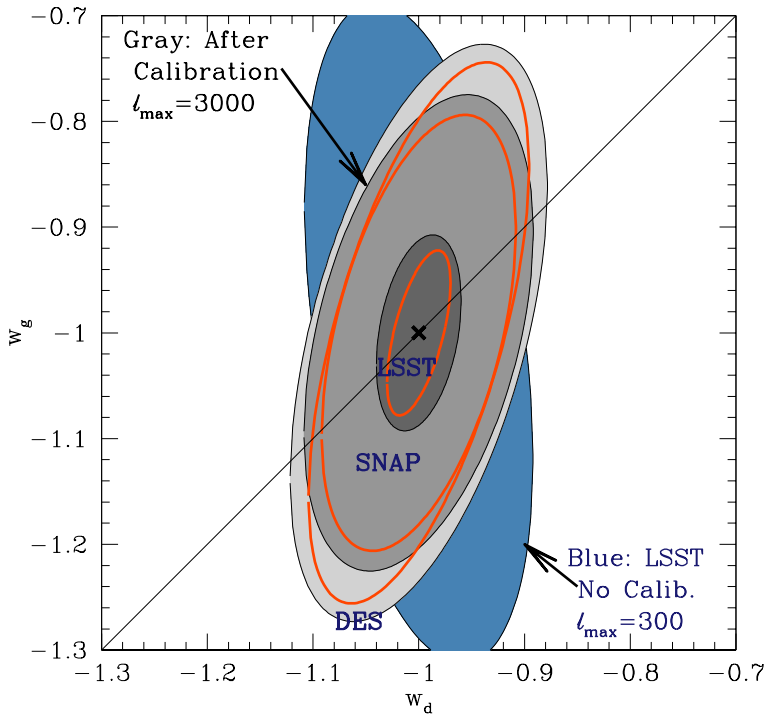

Figure 8: Constraint contours in the $w_{\rm d}$-$w_{\rm g}$ plane after accounting for and marginalizing over the uncertainty in halo structure. The *filled, grey* contours correspond to the 1σ contours in this plane from the LSST-like (innermost contour), SNAP-like (middle contour), and DES-like (outermost blue contour) experiments we consider. The *thick, solid* lines overlaying the contours represent the confidence contours in the limit of perfect knowledge of nonlinear structure formation as shown in the left panel of Fig. 4. The cross shows the fiducial model and the diagonal line running from the lower left to the upper right delineates $w_{\rm d}$=$w_{\rm g}$. As an explicit demonstration of the effect of excising small-scale information to eliminate potential bias, the *outermost, blue* contour shows the 1σ marginalized constraint in the $w_{\rm d}$-$w_{\rm g}$ plane from LSST if no information from multipoles greater than $\ell_{\rm max} = 300$ are utilized.

3.0 THE IMPACT OF PHOTOMETRIC REDSHIFT ERRORS ON WEAK LENSING STUDIES OF DARK ENERGY

3.1 INTRODUCTION

As seen in the last chapter, weak gravitational lensing of galaxies by large-scale structure is developing into a powerful cosmological probe [e.g., 11; 46; 56; 71; 86; 87; 99; 143; 165; 187]. Forthcoming imaging surveys such as the Dark Energy Survey (DES), the Large Synoptic Survey Telescope (LSST), the European Space Agency's Euclid, and the Joint Dark Energy Mission (JDEM) expect to exploit measurements of weak gravitational lensing of distant source galaxies as one of the most effective means to constrain the properties of the dark energy [e.g., 4; 43; 65; 70; 72; 76; 78; 84; 130; 148; 149; 173; 179; 180; 197; 198; 204]. The most stringent dark energy constraints can be achieved when source galaxies can be binned according to their redshifts, yielding a tomographic view of the lensing signal. Among the contributions to the dark energy error budget will be the error induced by the need to use approximate redshifts determined from photometric data [2; 8; 17; 19; 30; 38; 39; 41; 54; 83; 109; 135; 139; 161] because it is not possible to obtain spectroscopic redshifts for the large numbers of source galaxies needed to trace cosmic shear. Photometric redshifts are and will be calibrated by smaller samples of galaxies with spectroscopic redshifts. In this chapter, we study the influence of poorly-calibrated photometric redshifts for small subsets of the galaxies within imaging samples on dark energy constraints.

The influence of uncertain photometric redshifts (photo-z hereafter) on the dark energy program has been studied by a number of authors [13; 82; 100; 108; 118; 119; 178; 196; 203]. Studies of the requirements for photo-z accuracy have assumed relatively simple forms for the relationship between the inferred photo-z of a galaxy and its spectroscopic redshift, in par-

ticular that this is a Gaussian distribution with a redshift-dependent bias and scatter. The underlying assumption is that this distribution can be calibrated with an appropriate spectroscopic sample over the range of redshifts of interest. These studies indicate that roughly $N_{\rm spec} \sim 10^5$ spectroscopic redshifts are needed to render photo-z uncertainty a small contributor to the dark energy error budget, but any particular number depends upon the many details of each study. Broadening the description of the photo-z distribution to a multi-component Gaussian leads to slightly more demanding requirements on the spectroscopic calibration sample [118]. However, complexity or multi-modality of the photo-z distribution will not induce large systematic errors on dark energy parameters, provided that this complexity is known and that we have some ability to calibrate complex photo-z features using spectroscopic galaxy samples (a non-trivial assumption). That is not to say that dark energy constraints are insensitive to such complexity. Broad or multi-modal photo-z distributions will provide an effective limit to the redshift resolution of tomographic weak lensing and will degrade dark energy constraints. If this complexity can be diagnosed in spectroscopic samples, it may be treated by generalizing the modeling in Ma et al. [119] and Ma and Bernstein [118]. In approximate accordance with the prevailing nomenclature, we refer to the galaxies for which spectroscopic calibration of the photo-z distribution will be possible as the *core* photo-z distribution. Ma and Bernstein [118] studied multi-modal core photo-z distributions in some detail.

These studies assume that the spectroscopic samples that will be obtained will suffice to calibrate the photo-z's of all galaxies utilized in the weak lensing analysis. However, spectroscopic calibration samples may well be deficient in spectra of some subset of galaxies that otherwise may not be easily identified and removed from the imaging sample [for example, see 135, for a discussion]. Consequently, some fraction of galaxies in forthcoming imaging samples may not have photo-z's that are well calibrated spectroscopically and may have photo-z's that differ markedly from their true redshifts. Including such galaxies in weak lensing analyses would lead one to infer biased estimators of dark energy parameters. These systematic offsets in dark energy parameters may be considerable compared to statistical errors.

We refer to such subsets of galaxies that are not well calibrated by spectroscopic samples

and which have photo-z distributions that differ markedly from the photo-z distributions of the core galaxy samples as *catastrophic* photo-z outliers. Our chief aim in this study is to estimate the biases induced on dark energy estimators by catastrophic photo-z outliers for a variety of possible manifestations of catastrophic outliers, and to estimate the level at which such outliers must be controlled in order to mitigate dark energy biases.

We consider two broad classes of catastrophic outliers, differentiated by the breadth of their photo-z distributions. We emphasize that our definition of a catastrophic outlier is more inclusive than previous usage [compare to 13]. Outlier populations with photo-z's that are confined to a small range of highly-biased redshifts make up the class we refer to as *localized catastrophes*. As an example, such outliers may correspond to galaxy populations in which spectral features have been misidentified in broadband photometric observations; the prevailing usage of the term *catastrophic error* closely resembles our usage of the term *localized catastrophe*. The second class of outliers we consider, which we refer to as *uniform catastrophes*, have photometric redshifts that are relatively unconstrained. This class may more naturally be associated with a level of spectroscopic incompleteness yielding a population of imaged galaxies with little information on the reliability of their photometric redshifts.

We describe our modeling techniques in § 3.2. We detail our results on the potential importance of catastrophic photo-z outliers in § 3.3. This section includes a brief discussion of mitigation strategies in which we explore the possibility of eliminating subsets of galaxies in order to reduce biases at the cost of increased statistical errors. We discuss the implications of our results in § 3.4 and summarize our work in § 3.5.

We include in this study an appendix that may help in comparing published results on photometric redshift calibration requirements. All treatments of dark energy constraints from weak lensing rely on some approximate treatment of the growth of structure in the nonlinear regime. Several approaches are in common use [33; 117; 141; 157; 162; 164; 171] and additional parameters have been introduced to model baryonic processes [59; 155; 197]. In the main body of this chapter, we use the fitting form provided by Smith et al. [171]. We demonstrate in the appendix that estimates of photo-z calibration requirements depend upon the modeling of nonlinear power. Implementing the Smith et al. [171] relation for

nonlinear power results in significantly reduced photo-z calibration requirements as compared to previous results [e.g., 119] that employed the Peacock and Dodds [141] approximation.

3.2 METHODS

In this section we describe the methods used in our analysis. We begin in § **??** with a discussion of our treatment of photometric redshifts, including both the *core* photometric redshift distributions as well as *catastrophic* outliers. In § 3.2.2, we describe our weak lensing power spectrum observables. We describe cosmological parameter forecasting in § 3.2.3 and conclude with a description of our fiducial cosmology and representative surveys in § 3.2.4.

3.2.1 Photometric Redshift Distributions of Source Galaxies

We characterize the distribution of photometric redshifts through the probability of obtaining a photometric redshift $z^{\rm ph}$, given a galaxy with spectroscopic (or "true") redshift z, $P(z^{\rm ph}|z)$. The distribution of true redshifts of galaxies in a photometric bin labeled with index i is

$$n_i(z) = n(z) \int_{z_i^{\rm low}}^{z_i^{\rm high}} {\rm d}z^{ph}\ P(z^{\rm ph}|z), \tag{3.1}$$

where $n(z)$ is the number density of source galaxies per unit redshift z, $n_i(z)$ is the number density of sources per unit redshift that are assigned to the $i^{\rm th}$ photo-z bin, and $z_i^{\rm low}$ and $z_i^{\rm high}$ delineate the boundaries of the $i^{\rm th}$ tomographic bin.

We model the overall galaxy distribution via

$$n(z) \propto z^2\ \exp[-(z/z_0)^{1.2}], \tag{3.2}$$

where z_0 is determined by specifying the median redshift of the survey and the powers of redshift are representative of the distributions of observed high-redshift galaxies [136]. The normalization of the overall galaxy distribution is determined by the total number of galaxies per unit solid angle,

$$N^A = \int_0^\infty {\rm d}\ z\ n(z),$$

and we designate the number of galaxies per solid angle in any photo-z bin as

$$N_{\rm i}^A = \int_0^\infty \mathrm{d}\ z\ n_i(z).$$

3.2.1.1 The Core Photometric Redshift Distribution For the purposes of our study, we consider the *core* galaxy distribution to be comprised of galaxies with a *photometric* redshift distribution that will be well characterized through calibration with spectroscopically-observed galaxy samples. Studies using existing spectroscopic galaxy samples to predict the photo-z distributions of galaxies in future large-scale image surveys indicate that the core distributions may be complicated [e.g., 38; 83; 92]. A common simplifying assumption in the literature is that the photometric redshifts of galaxies in the core are distributed according to a Gaussian distribution with a redshift-dependent mean and variance [e.g., 118; 119],

$$P_{core}(z^{ph}|z) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{(z - z^{ph} - z^{bias})^2}{2\sigma_z^2}\right], \tag{3.3}$$

where both $\sigma_z(z)$ and $z^{\rm bias}(z)$ are functions of true redshift, z. The redshift-dependent mean and variance endow this form with sufficient flexibility to treat a wide variety of redshift distributions; however, this simple model does neglect complex features that may be present in the realized photometric redshift distributions of future surveys. We adopt this model because it is a published standard against which our results can be compared, and because the complexity of calibrating the core sample of photometric redshifts is not the primary aim of our work.

We compute the functions $\sigma_z(z)$ and $z^{\rm bias}(z)$ by linear interpolation between values tabulated at 31 redshift points spaced evenly between $z = 0$ and $z = 3$. This choice of binning allows for maximal degradation in dark energy constraints absent prior information about the photometric redshift distribution of source galaxies. We treat the bias and dispersion at each of these redshifts as free parameters in our forecasts, so that there are $2 \times 31 = 62$ free parameters describing the core photometric redshift distribution. For our fiducial model, we take $\sigma_z(z) = 0.05(1 + z)$ and $z^{\rm bias}(z) = 0$.

3.2.1.2 Catastrophic Photometric Redshift Outliers Forthcoming large imaging surveys will observe a tremendous number of galaxies. It is unlikely that accurate calibration of every class of photometric redshift distribution will be made, at least in part due to the limitations of obtaining reliable spectroscopic redshifts [e.g., 135] and observations of relatively rare objects. If either the uncalibrated objects follow the redshift distributions of the sample of calibrated photometric redshifts, or the uncalibrated objects can be identified from imaging data and removed from the sample, they will have a relatively benign impact on the dark energy aims of these surveys. In the former case, they present no systematic error because they follow the redshift distribution of the majority of galaxies, and in the latter case they can be removed from the imaging sample at a small cost in statistical uncertainty. Conversely, if a sample of uncalibrated source galaxies that does not follow the redshift distribution of the calibrated sources remains in the imaging data used for dark energy constraints, this could represent a significant additional systematic error. In approximate accordance with established nomenclature, we refer to subsets of galaxies that do not follow calibrated photometric redshift distributions *and* cannot be removed from imaging data as *catastrophic* photometric redshift outliers.

In practice, it is expected that catastrophic photometric redshift outliers will be present at some level in forthcoming imaging surveys. The prevalence of multi-modal features in the photo-z distributions of existing calibration samples is a clear illustration of the difficulty of determining galaxy redshifts from photometric colors [38; 39; 83; 139]. When a population of galaxies responsible for a non-trivial photometric redshift determination appears sufficiently often in spectroscopic samples, its associated photo-z error can be calibrated, perhaps leading to multi-modal features in the core distribution. However, there will inevitably be populations of galaxies with photo-z degeneracies that are sufficiently rare so as to evade spectroscopic sampling, the spectroscopic calibration of a truly representative sample will not be complete, and the removal of galaxies with troublesome redshifts from the imaging data will be imperfect. Each of these difficulties leads to a population of outlier galaxies, with distributions not described by the core photometric redshift model, that contributes a systematic error to dark energy parameter estimators.

To illustrate the distinction between catastrophic outliers and multi-modal features in the

core, consider the photo-z distribution illustrated in Figure 9. The bulk of the galaxies in this distribution (black diamonds) are scattered about the line $z = z^{\rm ph}$. This is a population of 400 galaxies drawn from the Gaussian distribution of Eq. (3.3). There are also two "islands" in the distribution. The appearance of these island contributions to the photo-z distribution is quite similar, but they are intended to represent photo-z errors of a qualitatively different nature, as discussed below. One island has $(z, z^{\rm ph})$ coordinates $(0.3, 3.7)$ and the other has $(2.0, 0.8)$. The island at $(z, z^{\rm ph}) = (0.3, 3.7)$, consisting of black squares, is a schematic representation of some subset of galaxies that give a known, calibrated, small probability of yielding a highly-biased photometric redshift. This is a component of a multi-modal core distribution and may either be calibrated with spectroscopy or removed from the sample. The island at $(z, z^{\rm ph}) = (2.0, 0.8)$, consisting of the red crosses, is a schematic representation of a catastrophic outlier population. These are a small subset of galaxies with true redshifts near $z \approx 2$, that yield strongly biased, but localized, photometric redshifts. Moreover, this is a population that is either not identified and calibrated in spectroscopic samples, or is incompletely removed from imaging data, so that this outlier contributes a systematic error to the dark energy error budget. This is the type of error that is our focus in this chapter. Finally, there is a population of galaxies that is localized near $z \approx 1$ and spread uniformly across $z^{\rm ph}$. These galaxies represent another extreme of catastrophic photo-z errors in that the redshifts may not be strongly biased, but they are poorly constrained and will contribute systematic errors for dark energy.

One cause for localized catastrophic redshift errors (such as the *red crosses* in Fig. 9) is the misidentification of a spectral feature in broadband photometric observations of galaxies over some range of true redshift. A specific example of this occurs when the Lyman-break is confused with the $4000A$-break. The effect on the photo-z distribution of a small portion of errors due to Lyman-$4000A$ confusion would look something like the small island of squares at $(z, z^{\rm ph}) = (0.3, 3.7)$ in Fig. 9 [13]. Confusion between the Lyman and $4000A$ breaks may occur often enough in spectroscopic samples to be calibrated and thus included as a secondary peak in the core distribution, but it is possible that there will be other small redshift windows where tertiary islands remain uncalibrated.

Throughout this chapter we adopt a simple model for the photo-z distributions of *localized*

catastrophes as Gaussians with spreads σ_{cat} centered away from the core at $z^{\rm ph}_{\rm cat}$,

$$P_{\rm cat}(z^{\rm ph}|z) = \frac{1}{\sqrt{2\pi}\sigma_{cat}} \exp\left[-\frac{(z^{\rm ph} - z^{\rm ph}_{\rm cat})^2}{2\sigma^2_{cat}}\right]. \tag{3.4}$$

The parameter $z^{\rm ph}_{\rm cat}$ specifies the location of the island in photometric redshift, and $\sigma_{\rm cat}$ gives the spread of the catastrophe in $z^{\rm ph}$. In the presence of a localized catastrophe the total photometric redshift distribution is

$$\begin{aligned} P_{\rm tot}\left(z^{ph}|z\right) &= \left[1 - \Xi(z)F_{cat}\right] P_{core}\left(z^{ph}|z\right) \\ &+ \Xi(z)F_{cat}P_{cat}\left(z^{ph}|z\right). \end{aligned} \tag{3.5}$$

The catastrophic error occurs over only a specified range of true redshifts, $z_{\rm cat} - \Delta z_{\rm cat}/2 < z < z_{\rm cat} + \Delta z_{\rm cat}/2$, as enforced by the function

$$\Xi(z) \equiv \Theta\left(\frac{\Delta z_{\rm cat}}{2} - |z - z_{\rm cat}|\right), \tag{3.6}$$

where $\Theta(x)$ is the Heaviside step function. The quantities $z^{\rm ph}_{\rm cat}$ (location of the local catastrophe in photometric redshift), $z_{\rm cat}$ (central value of the range of true redshifts over which the catastrophe occurs), $\Delta z_{\rm cat}$ (width of the range of true redshifts over which the catastrophic error is made), and $\sigma_{\rm cat}$ (spread in $z^{\rm ph}$ of the catastrophe) are four of the five parameters that specify the local catastrophe model. The fifth parameter, $F_{\rm cat}$, is the fraction of galaxies in the true redshift window set by $z_{\rm cat}$ and $\Delta z_{\rm cat}$ for which the catastrophic error occurs.

The term $\Xi(z)F_{\rm cat}$ removes the appropriate fraction of galaxies from the core distribution and ensures that $\int_0^\infty {\rm d}z^{\rm ph}\ P_{\rm tot}\left(z^{ph}|z\right) = 1$. As a concrete example, the catastrophic outliers represented by the red crosses in Fig. 9 are galaxies drawn from our model with $F_{\rm cat} = 0.03$, $z^{\rm ph}_{\rm cat} = 0.8$, $z_{\rm cat} = 2.0$, $\sigma_{\rm cat} = 0.1$, and $\Delta z_{\rm cat} = 0.1$. For the sake of pragmatism, we present results for localized catastrophes in interesting limits of this five-dimensional parameterization rather than an exhaustive exploration of these parameters.

Empirically, photometric redshift determination algorithms applied to extant calibration samples yield photometric redshift estimates that are relatively unconstrainted on some subsets of galaxies. For example, the photo-z distribution of galaxies in both CFHTLS [38] and COSMOS [83] possess such a feature within the range of error rates we explore

in this work. Unconstrained photometric redshifts represent a regime complementary to localized catastrophes. In this case, photometric redshifts may be obtained with nearly equal probability over a significant range of redshift. Such broad errors may occur when light from one galaxy is contaminated by light from another source nearby in angular separation but at a different redshift. It is natural to expect that such errors will occur most frequently near the peak of the observed galaxy number density $n(z)$.

Motivated by the presence of such errors, we also treat the extreme case of relatively unconstrained photometric redshifts by using a uniform distribution for $z^{\rm ph}$, over a symmetric window in true redshift centered on $z_{\rm cat}$ and spanning a width of $\Delta z_{\rm cat}$. We refer to this kind of error as a *uniform* catastrophe for simplicity. In the presence of a uniform catastrophe the total photometric redshift distribution is

$$\begin{aligned} P_{\rm tot}\left(z^{ph}|z\right) &= \left[1-\Xi(z)F_{\rm cat}\right]P_{core}\left(z^{ph}|z\right) \\ &+ \Xi(z)F_{\rm cat}/(z_{\rm max}-z_{\rm min}), \end{aligned} \tag{3.7}$$

where $z_{\rm min}$ and $z_{\rm max}$ delineate the photometric redshift range of the survey. In analogy to localized catastrophes, the function $\Xi(z)$ restricts the true redshift range over which flat catastrophes occur and F_{cat} specifies the fraction of galaxies in this true redshift window whose redshifts are catastrophically in error. Therefore, three parameters specify this simple model, namely $F_{\rm cat}$, $z_{\rm cat}$, and $\Delta z_{\rm cat}$. The uniform catastrophe represented by the red triangles in Fig. 9 is drawn from a model with $F_{\rm cat}=0.05$, $z_{\rm cat}=1.0$, and $\Delta z_{\rm cat}=0.2$.

3.2.2 Cosmic Shear Tomography

In this study, we consider constraints from weak gravitational lensing observables only. We split source galaxies into $N_{\rm TOM}$ photometric redshift bins and consider as our observables the $N_{\rm TOM}(N_{\rm TOM}+1)/2$ distinct number-weighted auto- and cross-power spectra of convergence among the source redshift bins. Unless otherwise stated, we bin source galaxies in equal intervals of redshift between $z^{\rm ph}=0$ and $z^{\rm ph}=3$ and take $N_{\rm TOM}=5$, resulting in 15 distinct observables. For this redshift range, five-bin tomography is a useful standard because this

binning scheme suffices to saturate dark energy constraints [119, we have verified that this remains so within the parameters of our study as well].

The galaxy number count in each tomographic bin N_i^A, the cross-spectra between bins i and j, $P_\kappa^{\rm ij}(\ell)$, and the number-weighed spectra $\mathcal{P}_\kappa^{\rm ij}(\ell)$ are related by

$$\mathcal{P}_\kappa^{\rm ij}(\ell) = N_i^A N_j^A P_\kappa^{\rm ij}(\ell) = \int_0^\infty dz \frac{W_i(z) W_j(z)}{H(z) D_A^2(z)} P_\delta(k = \ell/D_A, z). \tag{3.8}$$

In Eq. (3.8), ℓ is the multipole number, $H(z)$ is the Hubble expansion parameter, $D_A(z)$ is the angular diameter distance to redshift z, and $P_\delta(k, z)$ is the three-dimensional matter power spectrum. The lensing weight functions, $W_i(z)$, weight the cosmic shear signal according to the redshift distributions of galaxies within each tomographic bin and are defined as

$$W_i(z) = \frac{3}{2}\Omega_{\rm m} H_0^2 (1+z) D_A(z) \int_z^\infty dz' \frac{D_A(z, z')}{D_A(z')} n_i(z'), \tag{3.9}$$

where $D_A(z, z')$ is the angular diameter distance between redshifts z and z'.

3.2.3 Parameter Forecasting

We use the Fisher matrix formalism to study the constraining power of our weak lensing observables on dark energy parameters as well as to quantify the systematic errors on dark energy parameters that result from catastrophic photometric redshift errors. The Fisher matrix formalism is ubiquitous in cosmological parameter forecasting [useful references related to the present application include, 4; 13; 93; 103; 163; 181], so we simply quote relevant results here. The particular implementation we use closely mirrors that in [197] and Hearin and Zentner [63], to which we refer the reader for details.

The Fisher matrix is given by a sum over the observables. In the particular case of weak lensing power spectra, the spectra at different multipoles can be treated as independent and this sum can be written as

$$F_{\alpha\beta} = \sum_{\ell_{\rm min}}^{\ell_{\rm max}} (2\ell + 1) f_{\rm sky} \sum_{\rm A,B} \frac{\partial \mathcal{P}_{\rm A}}{\partial p_\alpha} [C^{-1}]_{\rm AB} \frac{\partial \mathcal{P}_{\rm B}}{\partial p_\beta} + F_{\alpha\beta}^{\rm P}, \tag{3.10}$$

where the $\mathcal{P}_{\rm A}$ are the set of observables indexed by a single label, $\mathbf{C}^{-1}$ is the inverse covariance matrix of these observables at fixed multipole, $[C^{-1}]_{\rm AB}$ are the components of the inverse of

the covariance matrix (we include the brackets for clarity), and p_α are the theoretical model parameters. We choose an indexing scheme in which lower-case Greek letters designate model parameters, upper-case Latin letters designate observables, and lower-case latin letters designate photometric redshift bins, and take the mapping between observable number and tomographic bin number to be $A = i(i-1)/2 + j$. Throughout this chapter we use $\ell_{min} = 2f_{\rm sky}^{-1/2}$, where $f_{\rm sky}$ is the fractional sky coverage of the weak lensing survey, and $\ell_{max} = 3000$ as a rough indication of the scale beyond which a number of weak lensing approximations break down [32; 44; 155; 165; 185; 192].

The covariance matrix of observables at each multipole is

$$C_{\rm AB}(\ell) = \bar{\mathcal{P}}_\kappa^{\rm ik}(\ell)\bar{\mathcal{P}}_\kappa^{\rm jl}(\ell) + \bar{\mathcal{P}}_\kappa^{\rm il}(\ell)\bar{\mathcal{P}}_\kappa^{\rm jk}(\ell) \tag{3.11}$$

where the indices i and j map onto A and k and l map onto B. The observed number-weighted power spectra, $\bar{\mathcal{P}}_\kappa^{\rm ij}(\ell)$, have contributions from signal and shot noise,

$$\bar{\mathcal{P}}_\kappa^{\rm ij}(\ell) = \mathcal{P}_\kappa^{\rm ij}(\ell) + N_i^A \delta_{ij} \langle \gamma_{\rm i}^2 \rangle, \tag{3.12}$$

where the quantity $\langle \gamma_{\rm i}^2 \rangle$ is the intrinsic source galaxy shape noise. We conform to recent convention and fix $\sqrt{\langle \gamma_{\rm i}^2 \rangle} = 0.2$, so that all deviations from this noise level are incorporated into an effective galaxy number density.

The Fisher matrix formalism provides an estimate of the parameter covariance near a fiducial point in the parameter space. One chooses fiducial values for the model parameters and estimates the error on parameter α from the inverse of the Fisher matrix at this point, $\sigma(p_\alpha) = [F^{-1}]_{\alpha\alpha}$. Within this formalism, statistically-independent prior information about the parameters is easily incorporated by simple matrix addition. The second term in Eq. (3.10) is the prior matrix. In our analysis, we assume *independent* prior constraints on cosmological parameters, so that the prior matrix reduces to a simple diagonal matrix, $F^P_{\alpha\beta} = \delta_{\alpha\beta}/\sigma_\alpha^2$, where $\delta_{\alpha\beta}$ is the Kronecker-δ symbol and σ_α is the prior 1-σ, Gaussian constraint on parameter p_α.

Given a systematic error that induces a specific shift in the observables, one can use the Fisher matrix to estimate the ensuing systematic error in model parameters. Using $\Delta\mathcal{P}_{\rm A}$ to denote the difference between the fiducial observables and the observables perturbed by

the presence of the systematic error, one will infer a set of parameters that is systematically offset from the true parameters by

$$\delta p_\alpha = \sum_\beta [F^{-1}]_{\alpha\beta} \sum_\ell (2\ell+1) f_{\rm sky} \sum_{\rm A,B} \Delta\mathcal{P}_{\rm A} [C^{-1}]_{\rm AB} \frac{\partial \mathcal{P}_{\rm B}}{\partial p_\beta}. \tag{3.13}$$

The primary results of our work are estimates of the systematic errors in dark energy parameters induced by catastrophic photometric redshift outliers. In related literature, the δp_α are often referred to as *biases*; however, we refer to them as *systematic errors* in order to avoid potential confusion with the biases in photometric redshifts.

3.2.4 Cosmological Model and Survey Characteristics

We assume a cosmological model specified by seven parameters. Three of these parameters describe the dark energy. These three parameters are the present energy density in units of the critical density, $\Omega_{\rm DE} = 0.76$, and two parameters, $w_0 = -1$ and $w_{\rm a} = 0$, that describe a linearly-evolving dark energy equation of state, $w(a) = w_0 + (1-a)w_{\rm a}$. The values specified for these parameters are those in our fiducial cosmological model. In models with a time-varying dark energy equation of state it is interesting to present results for the constraint on $w(a)$ at the scale factor at which it is most well constrained. The scale factor at which $w(a)$ can be best constrained is the *pivot* scale factor $a_{\rm p}$, and is related to the Fisher matrix components as

$$a_{\rm p} = 1 + \frac{[F^{-1}]_{w_0 w_{\rm a}}}{[F^{-1}]_{w_{\rm a} w_{\rm a}}}. \tag{3.14}$$

The pivot equation of state parameter is

$$w_{\rm p} \equiv w(a_{\rm p}) = w_0 + (1-a_{\rm p}) w_{\rm a} \tag{3.15}$$

and the error on $w_{\rm p}$ is

$$\sigma^2(w_{\rm p}) = [F^{-1}]_{w_0 w_0} - \frac{([F^{-1}]_{w_0 w_{\rm a}})^2}{[F^{-1}]_{w_{\rm a} w_{\rm a}}}. \tag{3.16}$$

The dark energy task force quantifies the constraining power of forthcoming surveys according to a figure of merit that reflects the areas of the confidence ellipses in the w_0-$w_{\rm a}$ plane. In particular, the task force quotes values for the combination $\mathcal{F} \equiv [\sigma(w_{\rm a}) \times \sigma(w_{\rm p})]^{-1}$ [4].

The other cosmological parameters we consider and the fiducial values they assume in our modeling are: the non-relativistic matter density $\omega_{\rm m} \equiv \Omega_{\rm m} h^2 = 0.13$; the baryon density $\omega_{\rm b} = \Omega_{\rm b} h^2 = 0.0223$; the amplitude of the primordial curvature fluctuations $\Delta^2_{\mathcal{R}} = 2.1 \times 10^{-9}$ (though in practice we vary $\ln \Delta^2_{\mathcal{R}}$ when computing derivatives of this parameter) evaluated at the pivot scale $k_{\rm p} = 0.05$ Mpc^{-1}; and the power-law index of the spectrum of primordial density fluctuations $n_{\rm s} = 0.96$. We adopt relatively conservative priors of $\sigma(\omega_{\rm m}) = 0.007$, $\sigma(\omega_{\rm b}) = 10^{-3}$, $\sigma(\ln \Delta^2_{\mathcal{R}}) = 0.1$, and $\sigma(n_{\rm s}) = 0.04$, each of which is comparable to contemporary, marginalized constraints on these parameters [101].

In principle, it is relatively straightforward to scale parameter forecasts from one experiment to another [e.g., 13; 119]; however, in the interest of simplicity, we present explicit results for three specific experimental configurations that span the range of observations expected of forthcoming instruments.

The Dark Energy Survey is the most near-term survey that we consider [1]. We model a DES-like survey by assuming a fractional sky coverage of $f_{\rm sky} = 0.12$ and a surface density of imaged galaxies of $N^A = 15/\text{arcmin}^2$. Second, we consider a narrow, deep imaging survey similar to a Supernova Acceleration Probe-like implementation of a JDEM[2,3]. We refer to this second type of survey as DEEP and model it with $f_{\rm sky} = 0.05$ and $N^A = 100/\text{arcmin}^2$. Lastly, motivated by a future ground-based imaging survey such as may be carried out by the LSST[4] [3], or a space-based mission such as the European Space Agency's Euclid[5] [150], we consider a survey with very wide sky coverage taking $f_{\rm sky} = 0.5$ and $N^A = 30/\text{arcmin}^2$. We refer to this class of survey as WIDE. We assume that the median galaxy redshift in the WIDE and DEEP surveys is $z_{\rm med} = 1.0$ and that the median galaxy redshift in the DES-like survey is $z_{\rm med} = 0.7$. In all cases, we follow recent convention by taking the shape noise to be $\sqrt{\langle \gamma^2 \rangle} = 0.2$, subsuming additional noise contributions into an effective galaxy number density. Table 1 summarizes our assumed survey properties.

[1] http://www.darkenergysurvey.org
[2] http://universe.nasa.gov/program/probes/jdem
[3] http://snap.lbl.gov/
[4] http://www.lsst.org
[5] http://sci.esa.int/euclid

3.3 RESULTS: SYSTEMATIC ERRORS ON THE DARK ENERGY EQUATION OF STATE

In this section, we present the results of our study of catastrophic photometric redshift outliers. We begin with the baseline constraints on the dark energy equation of state parameters in the limit of perfect knowledge of the source galaxy photometric redshift distribution in § 3.3.1. We continue in a sequence of increasing complexity. We quantify the influence of catastrophic photometric redshift errors in the limit of perfect knowledge of the core photometric redshift distribution in § 3.3.2. We present results on the influence of catastrophic photometric redshift errors in the more realistic case of imperfect knowledge of the core distribution in § 3.3.3. We explore the prospect of excising galaxies based on their photometric redshifts as a simple, first-line defense against systematic errors induced by catastrophic photometric redshift errors in § 3.3.4.

3.3.1 Baseline Constraints

We begin our results section by stating our forecasts for dark energy constraints in the limit of perfect knowledge of the photometric redshift distribution. With little uncertainty in photometric redshift distributions, the statistical limits of forthcoming survey instruments would allow for constraints on the dark energy equation of state at the level of a few percent, as summarized in Table 1. We emphasize here that the limit of perfect knowledge of the photo-z distributions is not the assumption that photometric redshifts are precisely equal to the true redshifts of the source galaxies. Rather, the assumption is that there are no catastrophic errors, and that the photometric redshift distribution is described by the Gaussian in § 3.2.1.1 such that all 62 parameters used to specify the Gaussian distribution are known precisely.

3.3.2 Systematic Errors in The Limit of Perfect Core Knowledge

In this section, we present results for systematic photometric redshift errors in the limit of perfect knowledge of the *core* distribution of photometric redshifts. This amounts to the

assumption of prior knowledge of the 31 dispersion $[\sigma_z(z)]$ and 31 bias $[z_{\rm bias}(z)]$ parameters defined in § 3.2.1.1 to a level of $\lesssim 10^{-3}$, which could be achieved with a *representative* sample of $\gtrsim 4 \times 10^5$ spectroscopic redshifts distributed in redshift in a manner similar to those in the imaging survey [see 118; 119, and the discussion in the appendix of this dissertation]. This is a simple case to begin with as it allows exploration of the influence of catastrophic redshift errors over a range of the catastrophic photo-z parameter space without the additional complications associated with redshift-dependent priors on the core photo-z distribution. This is the limit explored by Bernstein and Huterer [13].

3.3.2.1 Uniform Catastrophes First, we address systematic errors induced on dark energy parameters by a small population of *uniform* catastrophes. Uniform catastrophes are cases in which some small population of galaxies with true redshifts in the range $(z_{\rm cat} - \Delta z_{\rm cat}/2 < z < z_{\rm cat} + \Delta z_{\rm cat}/2$ yield photometric redshift estimates that are distributed broadly in $z^{\rm ph}$. This class of error differs from the conventional use of the term *catastrophic error* and may more naturally be interpreted as a tolerance on spectroscopic incompleteness.

For simplicity, we take the central redshift of the uniform catastrophe to be $z_{\rm cat} = z_{\rm med}$, and determine systematic errors as a function of $\Delta z_{\rm cat}$, the width of the range of redshifts over which such errors occur, and $F_{\rm cat}$, the fraction of galaxies in this range of true redshift that correspond to this type of catastrophic error. We refer the reader to Eq. (3.6) and Eq. (3.7) for the expressions that formally define these parameters. While we vary these parameters independently, they are both related to the total number density of sources with redshifts that are catastrophically in error,

$$N^A_{\rm cat} = F_{cat} \int_{z_{\rm cat} - \frac{\Delta z_{\rm cat}}{2}}^{z_{\rm cat} + \frac{\Delta z_{\rm cat}}{2}} {\rm d}z' \; n(z'), \tag{3.17}$$

where $n(z)$ is the overall, true redshift distribution of galaxies. We should expect systematic errors to increase with both $F_{\rm cat}$ and $\Delta z_{\rm cat}$ because higher values of either parameter result in a greater total number of catastrophic errors in the outlier population.

In Figure 10 we have quantified the systematic errors induced by uniform catastrophic errors as a function of the parameters of our simple model. The curves in Fig. 10 are contours of constant systematic error on dark energy parameters (for example, $|\delta(w_0)|$ for

w_0) expressed in units of the statistical error ($\sigma(w_0)$ for w_0) at points in the $\Delta z_{\rm cat}$-$F_{\rm cat}$ plane. For each of the DES, Wide, and Deep surveys, the *solid* curves trace systematic errors in dark energy that are three times the statistical errors, while the *dashed* curves trace systematic errors that are 1/3 of the statistical error. For each of the surveys depicted in Fig. 10, the region of catastrophic parameter space that is bracketed by the solid and dashed curves labeled with the corresponding survey name corresponds to outliers that produce systematic errors which are comparable to statistical errors. Systematic errors are relatively small compared to statistical errors in the regions below the dashed curves. Each curve plotted in Fig. 10 has been generated with a fixed value of $z_{\rm cat} \equiv z_{\rm med}$; for the Wide and Deep surveys $z_{\rm med} = 1$, so, e.g., when $\Delta z_{\rm cat} = 2$ the uniform catastrophes are made over the true redshift range $0 < z < 2$; for DES $z_{\rm med} = 0.7$, so in Fig. 10 once $\Delta z_{\rm cat} > 1.4$ the true redshift window over which catastrophes are made only increases at the high-redshift boundary.

For each of the contours of constant systematic error in Fig. 10, $F_{\rm cat}$ decreases with increasing $\Delta z_{\rm cat}$. This is simply because increasing the redshift range over which the catastrophic errors are being made ($\Delta z_{\rm cat}$) leads to an increased total number of catastrophic errors, resulting in the survey's decreased tolerance to rate at which the error can be made ($F_{\rm cat}$). Alternatively, the total number of catastrophic errors in an outlier population is given by the integral in Eq. (3.17), and the contours of constant systematic error roughly trace constant values of $N^A_{\rm cat}$. The contours flatten considerably for errors that occur over a redshift range $\Delta z_{\rm cat} \gtrsim 0.4$ because there will be comparably few imaged sources with true redshifts near $z \sim 0$ or with $z \gtrsim 1.5$.

This treatment of a uniform $z^{\rm ph}$ catastrophe may appear somewhat contrived but it gives insight into a few basic results that are important to recognize. It is clear that the utility of forthcoming shear surveys to constrain dark energy is sensitive to a fractionally small population of galaxies that may yield poorly-determined photo-z estimates. If the error is only relevant to galaxies that are relatively isolated in narrow regions of true redshift, for example with $\Delta z_{\rm cat} \lesssim 0.1$, then error rates as high as $F_{\rm cat} \sim 1\%$ in this region of true redshift are tolerable. This is simply because errors that occur with a fixed rate over a small redshift range result in a small total number of catastrophic outliers to corrupt the weak

lensing tomography. On the contrary, if such an error occurs for a subset of galaxies with true redshifts in an interval of width $\Delta z_{\rm cat} \gtrsim 0.1$, then the error rate per galaxy must be significantly lower than $F_{\rm cat} \lesssim 0.01$ in order to render the systematic errors on dark energy equation of state parameters small.

The limit of $\Delta z_{\rm cat} \gg 0.1$ is interesting to consider. This may correspond to the case of a small fraction of galaxies that yield very poorly-constrained photometric redshifts over a broad range of true redshifts and that otherwise cannot be identified and removed from the imaging survey. In this case, the systematic error from catastrophic photometric redshifts becomes a considerable portion of the dark energy error budget at a rate of only $F_{\rm cat} \sim 10^{-3}$. Reducing the systematic error due to such an outlier population to a negligible level requires reducing the occurrence of such an outlier population to $F_{\rm cat} \lesssim 4 \times 10^{-4}$. Strictly speaking, Fig. 10 corresponds to errors that occur when the true redshift band over which the uniform catastrophe occurs is centered on $z_{\rm cat} = z_{\rm med}$, but for the $\Delta z_{\rm cat} \gtrsim 1$ limit, similar results hold for a wide range of $z_{\rm cat}$ near unity, so this result is of some general relevance to photometric redshift calibration studies.

3.3.2.2 Localized Catastrophes: Details Localized catastrophes correspond to the case where a small fraction of galaxies near some true redshift $z_{\rm cat}$ yield photometric redshifts that are narrowly distributed about a biased value $z_{\rm cat}^{\rm ph}$ that is very different from the true redshift $z_{\rm cat}$. Such errors could arise due to incomplete calibration by spectroscopic surveys or from difficulty in removing troublesome galaxies from the imaged galaxy sample. A known example of such an error occurs when photo-z algorithms confuse the $4000A$ break with the Lyman break, but other isolated islands of biased $z^{\rm ph}$ persist in contemporary photo-z algorithms (see, e.g. Coupon et al. 38, Ilbert et al. 83) and may be relevant to forthcoming imaging surveys.

The class of localized photo-z catastrophes is more complex than the uniform case because there are more relevant parameters needed to specify the manner in which a localized outlier population is distributed in $z^{\rm ph}$. Our toy model requires five parameters (see § 3.2.1.2 and Fig. 9 for an illustration). Two are the central value of the true redshift over which this error is operative ($z_{\rm cat}$) and the width of the true redshift range over which this error is

operative ($\Delta z_{\rm cat}$). Like in the uniform case, some fraction $F_{\rm cat}$ of galaxies with true redshifts in the interval $z_{\rm cat} - \Delta z_{\rm cat}/2 < z < z_{\rm cat} + \Delta z_{\rm cat}/2$ are catastrophically in error. The final two parameters specify the biased distribution of photometric redshifts that these galaxies are assigned. These are the (systematically erroneous) value of the photometric redshift $z_{\rm cat}^{\rm ph}$, and the dispersion in the catastrophic photometric redshift distribution $\sigma_{\rm cat}$ about $z_{\rm cat}^{\rm ph}$.

We make an effort to remain agnostic about the classes of photo-z errors that may be realized in future imaging data. However, a complete mapping of even the simple parameter space we have specified for catastrophic photo-z's would require a lengthy discussion, so we explore useful limits of the model parameters in order to distill our results into a small number of points. We are particularly interested in the limit where the source galaxies are placed in a narrow range of biased photo-z ($\sigma_{\rm cat} \ll 0.3$ or so) because the limit of large dispersion in the catastrophic photometric redshift population is similar to the *uniform* catastrophe of the previous section.

We first isolate the sinister regions in the space of $z_{\rm cat}$-$z_{\rm cat}^{\rm ph}$ that lead to the most destructive systematic errors in dark energy parameters. At a set of points in the parameter space of $(z_{\rm cat}, z_{\rm cat}^{\rm ph})$, we have calculated the systematic error induced in w_0 and $w_{\rm a}$ by distributing some fraction $F_{\rm cat}$ of the galaxies with true redshifts near $z_{\rm cat}$ in photometric redshifts centered around some $z_{\rm cat}^{\rm ph}$ that is generally very different from $z_{\rm cat}$. We sample a range of values of true redshifts from $z_{\rm cat} = 0.05$ to $z_{\rm cat} = 2.95$, evenly spaced in redshift intervals of $\delta z = 0.1$ and likewise for the photometric redshifts, $z_{\rm cat}^{\rm ph}$. In the interest of simplicity, we fix the remaining parameters of our catastrophic photo-z model to $F_{\rm cat} = 0.05$, $\Delta z_{\rm cat} = 0.05$, and $\sigma_{\rm cat} = 0.01$ to isolate the dependence of the parameter bias upon the location of the catastrophe.

It is important to note explicitly that we present results here at a fixed error fraction $F_{\rm cat}$, and a fixed true redshift window width and $\Delta z_{\rm cat}$. However, even with these parameters fixed the absolute number of errors varies with $z_{\rm cat}$ according to Eq. (3.17), which is roughly $N_{\rm cat}^A \sim n(z_{\rm cat})\Delta z_{\rm cat} F_{\rm cat}$ for sufficiently small $\Delta z_{\rm cat}$, along the lines of the analogous discussion for the uniform catastrophe in § 3.3.2.1. The aim of this calculation is to map out the relative importance of making errors at a *fixed rate per galaxy* as a function of the true and photometric redshifts of the outliers.

The results of this exercise are depicted in Figure 11. In each column of Fig. 11 there are two panels, corresponding to the systematic errors in w_0 and w_a, for each representative experiment. The horizontal axes show values of $z_{\rm cat}$ and the vertical axes show values of $z_{\rm cat}^{\rm ph}$. The systematic error is represented on the grid of $(z_{\rm cat}, z_{\rm cat}^{\rm ph})$ by the color in each of the cells. In discussing the results of this exercise, we find the terminology of [13] to be a useful, descriptive shorthand. We will refer to the tomographic bin that contains the $z_{\rm cat}$ value of an outlier as the *Source Bin* of that catastrophic photo-z population. We call the bin containing $z_{\rm cat}^{\rm ph}$ its *Target Bin*. This is because galaxies with true redshifts near $z_{\rm cat}$ are erroneously placed in the Target Bin containing the redshift $z_{\rm cat}^{\rm ph}$. Our sampling guarantees that no localized outlier straddles a tomographic bin boundary so there are always unique Source and Target bins. Outlier populations that straddle a boundary dividing two tomographic bins can be substantially more severe than those that do not because such an outlier simultaneously contaminates multiple Target bins. We have chosen to ignore such outlier populations for simplicity, but such outliers can be modeled by two catastrophic outlier populations, one for each affected target bin. We will return to the issue of tomographic binning and straddling outlier populations below.

The prominent block-like features in Fig. 11 reflect the tomographic redshift bins used in our analysis. The tomographic bins of the source and target galaxies largely determine both the magnitude and sign of the induced systematic error in dark energy parameters. This gives rise to features that reflect the structure of the photometric redshift binning in the $(z_{\rm cat},z_{\rm cat}^{\rm ph})$ plane. Indeed, for fixed Target and Source Bins, the specific value of the target redshift, $z_{\rm cat}^{\rm ph}$, within the target photometric redshift bin has little influence on the severity of the systematic error. However, small steps in $z_{\rm cat}^{\rm ph}$ can lead to large changes in systematic error when the boundary dividing two tomographic bins is crossed.

Varying the location in true redshift, $z_{\rm cat}$, leads to somewhat more significant changes in dark energy systematic error. Changing $z_{\rm cat}$ within fixed Source and Target Bins can result in up to a factor of two difference in systematic errors. Two factors primarily determine the severity of the systematic error as a function of the true redshift of the galaxies, $z_{\rm cat}$. The primary factor stems from the fact that a fixed fractional error rate ($F_{\rm cat}$) corresponds to a different absolute number of errors $N_{\rm cat}^A$ as a function of redshift, $z_{\rm cat}$. This is reflected in

Eq. (3.17). The number of errors $N_{\rm cat}^A$ will be relatively large in a region near the median redshift of the survey, where the number of source galaxies per unit redshift, $n(z)$, is largest. There are relatively few galaxies at low and high redshift, so for a fixed error rate, outlier populations with low or high true redshifts contribute a relatively small absolute number of galaxies with highly-biased redshifts.

Secondly, an outlier population naturally results in a more severe systematic error the more the photometric redshift is biased away from the true galaxy redshift. Consider the region of catastrophic parameter space near $(z_{\rm cat}, z_{\rm cat}^{\rm ph}) = (1.5, 2.7)$ in either color plot for DES. Catastrophes in this region of parameter space correspond to outlier populations whose Source bin is the third tomographic bin and Target bin is the fifth tomographic bin. Outliers in this region of parameter space are assigned photo-z's that are significantly too high. Decreasing the value of $z_{\rm cat}$ (the "source" of the error) increases the distance between core and outlier populations, thereby increasing the systematic error. This behavior contributes significantly to the systematic error gradient near $(z_{\rm cat}, z_{\rm cat}^{\rm ph}) = (1.5, 2.7)$ for DES in Fig. 11.

Each of the three representative experiments that we consider has two distinct "hot spots" in Fig. 11 that correspond to the most severe types of error given a fixed error rate per galaxy, $F_{\rm cat}$. A common feature of all these hot spots is their $z_{\rm cat}$ location. Each of the hot spots lies at a $z_{\rm cat}$ slightly beyond the median survey redshift. This is sensible because for a fixed error rate, the absolute number of catastrophic errors is greatest when they are made at the peak in the overall galaxy distribution, that is near $z_{\rm cat} = z_{\rm med}$. The most damaging systematic errors occur when the galaxies are shifted to either very low or very high photometric redshifts, when the target redshift, $z_{\rm cat}^{\rm ph}$, is very different from the source redshift, $z_{\rm cat}$, because the galaxies in error are then placed at distances significantly different from their true redshifts. For our WIDE and DEEP surveys the largest systematic errors tend to occur for galaxies shifted from a source redshift $z_{\rm cat}$ near $z_{\rm med} = 1$ to very low photometric redshifts.

The pattern of the DES catastrophic photo-z "hot spots" differs from that of the WIDE or DEEP surveys. Outliers with large values of $z_{\rm cat}^{\rm ph}$, that is those with a target in the fourth or fifth tomographic bin, are relatively more severe for DES. This is driven by the (assumed) comparably low redshift extent of imaged sources for a DES-like survey (with median redshift

$z_{\rm med} = 0.7$) This renders a contamination that extends to high redshift more disruptive due to the small population of galaxies with truly high redshifts. Though less striking, it is also evident in Fig. 11 that the DEEP survey is somewhat more sensitive than the WIDE survey to contamination of its fourth and fifth tomographic bins. The differences here are likewise driven by different survey depths and sky coverages. A deeper, but narrower survey (a JDEM perhaps) is relatively more sensitive to small-scale fluctuations induced by structure at high-redshift, so disruptions to the higher tomographic bins are more statistically significant for DEEP than for WIDE.

Finally, we return to the issue of tomographic binning with respect to the systematic errors in Fig. 11. We noted above that systematic errors in dark energy parameters can become markedly worse when the biased photometric redshifts ($z^{\rm ph}_{\rm cat}$) distribute galaxies across the boundary of a photometric redshift bin. The reason is because two sets of observables, namely the auto and cross spectra associated with the two target photometric redshift bins, become corrupted by the catastrophic photometric redshift error. The implication is that the level of systematic error induced by a localized catastrophic error is quite sensitive to photometric redshift binning. This is contrary to the statistical errors, which are insensitive to binning more finely than $N_{\rm TOM} \approx 5$ over the range $0 < z < 3$ [119].

Indeed this is the case. The general pattern shown in Fig. 11 is physically quite sensible and is robust to binning. However, in the case of localized catastrophes, binning more finely may reduce the absolute amplitude of systematic errors if the catastrophes do not occur near the edge of a photometric redshift bin. This is because smaller tomographic bins result in a smaller fraction of source galaxies that belong to a contaminated bin. This may be useful because even in the absence of significant prior indications of a localized catastrophe, re-analyzing the data with different photometric redshift binning schemes may reveal potential local catastrophes. In the least, it should be a useful strategy to choose photometric redshift bins such that suspect regions of $z^{\rm ph}$, where localized catastrophes may be anticipated, are contained in individual bins.

3.3.2.3 Localized Catastrophes: Summary A succinct distillation of the dominant effects that determine the structure of Fig. 11 is as follows. The systematic error induced by

a localized catastrophe will be most severe when:

1. $z_{\rm cat} \approx z_{\rm med}$, which maximizes the total number of outliers;
2. the distance between $z_{\rm cat}$ and $z_{\rm cat}^{\rm ph}$ is significant;
3. and when $z_{\rm cat}^{\rm ph}$ is such that the photometric redshift bins contain a fractionally large contaminant (in practice, high and low redshift extremes).

The details governing the magnitude of systematic errors generated by different regions of catastrophic error parameter space can be complicated. In general, these details depend on the relative statistical weights of the affected redshift bins, as well as the characteristics of the survey.

In isolation, Fig. 11 is useful in identifying the redshift errors that most seriously compromise dark energy constraints. A shortcoming of Fig. 11 is that we have assumed catastrophic errors that occur at a fixed rate of $F_{\rm cat} = 0.05$ and are active only over a range $\Delta z_{\rm cat} = 0.05$. The systematic errors induced on cosmological parameters scale approximately with the total number of catastrophes, $N_{\rm cat}^A$, in Eq. (3.17). In practice, scaling the systematic errors to new values of $\Delta z_{\rm cat}$ can be enacted over an interesting range of the parameter space by approximating $N_{\rm cat}^A \approx F_{\rm cat}\ n(z_{\rm cat})\ \Delta z_{\rm cat}$.

Figure 12 demonstrates the validity of scaling systematic error by the total number of errors, $N_{\rm cat}^A$, for three example localized catastrophes. Together, Fig. 11 and Fig. 12 provide a blueprint for estimating the systematic error induced by a wide range of localized catastrophes. One first reads off the systematic error level from Fig. 11 for the grid point of interest. For definiteness, suppose this systematic error in either of w_0 or $w_{\rm a}$ is δ. Provided that $\Delta z_{\rm cat}$ is small, one can approximate the systematic error induced by a different effective value of $\Delta z_{\rm cat}$ or $F_{\rm cat}$ (call it δ') by scaling δ in proportion to $N_{\rm cat}^A$ [Eq. (3.17)],

$$\delta' \approx \delta \times (F'_{\rm cat}/0.05) \times (\Delta z'_{\rm cat}/0.05)\,. \tag{3.18}$$

3.3.3 Catastrophic Redshift Errors with Core Uncertainty

In § 3.3.2, we presented results on the influence of catastrophic, uncalibrated photometric redshift errors on the systematic error budget for dark energy parameters w_0 and $w_{\rm a}$. In

that section, we assumed that the bulk of photometric redshifts had been well characterized by spectroscopy. In the nomenclature of this thesis and other papers, we assumed the limit in which the *core* of the photometric redshift distribution is calibrated so that its uncertainty does not contribute to the dark energy error budget. We developed guidance on how to optimally focus photo-z calibration efforts and identified the most severe types of catastrophes. In this section, we drop the assumption of arbitrarily precise calibration of the core populations of photometric redshifts. Our goal is to assess the relative importance of calibrating the core photometric redshift distribution compared to eliminating catastrophic errors.

We assume that the core photometric redshift distribution is specified by a Gaussian with redshift-dependent mean and dispersion. Following Ma et al. [119], we specify the unknown mean and dispersion at 31 points spaced evenly in redshift from $z = 0$ to $z = 3$ and allow for uncertainty in these parameters. In the interest of simplicity, we consider a one-parameter family for the prior knowledge about the core photometric redshifts that may be provided by a spectroscopic calibration sample. We do this by assuming a representative population of $N_{\rm spec}$ galaxies with spectroscopic redshifts, distributed evenly in redshift from $z = 0$ to $z = 3$, which can be used to calibrate the core photometric redshift distribution.

We implement core calibration by introducing priors on the values of the dispersion and bias at the ith point in redshift. These priors are

$$\Delta\sigma_z^i = \sigma_z^i \sqrt{\frac{1}{2N_{\rm spec}^{\rm i}}} \tag{3.19}$$

$$\Delta z_{\rm bias}^i = \frac{\sigma_z^i}{\sqrt{N_{\rm spec}^{\rm i}}} \tag{3.20}$$

where $z_{\rm bias}^i$ is the bias at the ith point in the tabulated core distribution, σ_z^i is the dispersion at this redshift, and $N_{\rm spec}^{\rm i}$ is the number of spectroscopic galaxies in each of the 31 bins of width $\delta z = 0.1$ used to calibrate the core photo-z redshift distribution. This prior model is certainly simplistic. For example, in our analysis we have chosen for the sake of simplicity to set all of the $N_{\rm spec}^{\rm i}$ equal to each other, so that our implementation assumes that calibrating spectra are sampled equally in redshift, whereas in reality we will have much looser constraints on sources at high redshift than those at low redshift. Moreover, both core calibration *and* the

ability to identify catastrophic outliers improve with larger spectroscopic samples. However, we consider these issues independently in the interest of completeness because the details of how a realistic calibration program may proceed remain uncertain.

Figure 13 is a contour plot depicting the systematic errors in w_0 and w_a induced by the worst-case-scenario catastrophes determined in § 3.3.2. The prior core knowledge is specified by $N_{\rm spec}$, which runs along the horizontal axis. The error rate, $F_{\rm cat}$, runs along the vertical axis. For uniform catastrophes the worst case outliers span the true redshift range of the survey. For localized catastrophes the most sinister outliers lie at the points of maximum systematic error in Fig. 11. The dashed (solid) curves are lines of constant systematic error at a level of one-third (three times) the statistical error on each parameter. Clearly then, systematic errors are dominant above the solid curves and become unimportant well below the dashed curves. Near the region bounded between the solid and dashed curves, statistical and systematic errors are comparable.

Several aspects of Fig. 13 are worthy of note. The contours all become very flat at large $N_{\rm spec}$. This is the limit in which the core photo-z distribution is calibrated sufficiently well that it no longer contributes to the error budget of w_0 and w_a [e.g., 118; 119]. This corresponds to the limit of perfect knowledge of the core photo-z distribution, and accordingly, the systematic errors asymptote to those quoted in § 3.3.2 at large $N_{\rm spec}$.

For a fixed level of systematic error, experiments generally become less tolerant of catastrophic outliers as $N_{\rm spec}$ increases. This behavior is reflected in the negative slope at the low $N_{\rm spec}$-end of the contours of constant systematic error in Fig. 13. This is an explicit manifestation of the competition between calibration of the "core" population of photometric redshifts and the ability to diagnose and eliminate a sub-dominant, poorly-understood "catastrophic" outlier population. The reason for this is simply that *systematic* errors must be better understood for samples with smaller *statistical* uncertainty. If the statistical errors in the measurement are intrinsically large, as they would be in the limit of poorly-calibrated photo-z's for the majority of the imaging sample, then high-rates of catastrophic outliers are tolerable because the systematic they contribute is not large compared to the statistical error induced by a poorly constrained core distribution.

When the core distribution is very well calibrated, most obvious at $N_{\rm spec} \gtrsim 10^5$ for

uniform catastrophes in DES, the contours of constant systematic error transition to slightly positive slope. This occurs when the core distribution has been sufficiently well calibrated that degeneracies between the photometric redshift parameters of the core distribution and cosmological parameters are no longer significant. Calibrating beyond the level required to break degeneracies between cosmology and the core photo-z parameters results in a slight reduction in systematic errors on cosmological parameters. This is a specific manifestation of the general result that improving priors can only lead to a net reduction in the systematic errors of inferred parameters, a result discussed in considerable detail in Bernstein and Huterer [13]. Clearly, the reduction in systematic error at very large $N_{\rm spec}$ is not significant in the cases of interest here.

To illustrate the competition between core calibration and the removal of outliers, consider some explicit examples. In the case of the uniform catastrophe, our Deep (Wide) survey can tolerate catastrophic errors at a rate $F_{\rm cat} > 1\%$ if the core calibration is worse than the statistical equivalent of $N_{\rm spec} \lesssim 3 \times 10^4$ ($N_{\rm spec} \lesssim 6 \times 10^4$). For both surveys, even the worst-case, localized catastrophes can occur at a rate of $F_{\rm cat} > 1\%$ if $N_{\rm spec} \lesssim 10^4$. Of course, the worst-case localized systematic errors are more subtle to interpret, as we have assumed they are only actively affecting galaxies over a range of true redshifts with width $\Delta z_{\rm cat} = 0.1$; however, the magnitude of the induced systematic errors produced by localized catastrophes active over different redshift ranges scales in proportion to $N^A_{\rm cat} \sim n(z_{\rm cat})\Delta z_{\rm cat} F_{\rm cat}$, as illustrated in Fig. 12. Detailed results are complex, but two simple conclusions are clear:

1. Limiting uniform catastrophic error rates to $F_{\rm cat} \lesssim 4 \times 10^{-4}$ ($F_{\rm cat} \lesssim 2 \times 10^{-4}$) for DES and DEEP (WIDE) will render them unimportant.

2. Limiting individual localized catastrophic error rates to $F_{\rm cat}(\Delta z_{\rm cat}/0.1) \lesssim 10^{-3}$ will render them unimportant for each experiment.

In practice, some amount of uncertainty in the calibration of the core distribution is inevitable, so error rates higher by a factor of a few may be tolerable, but in detail this will depend upon the nature of the error and the properties of the core sample of well-calibrated photometric redshifts. Fig. 11 and Fig. 13 contain the information necessary to diagnose the systematic error for a variety of idealized, but interesting cases.

3.3.4 Mitigating Systematic Errors by Sacrificing Statistics

In § 3.3.2 and § 3.3.3, we estimated the systematic errors that could be induced by two broad families of catastrophic photometric redshift error, remaining relatively agnostic about the source of the error. We found generally that error rates must be kept to levels below $F_{\rm cat} \sim 10^{-3}$, or one of a thousand imaged galaxies with large, uncalibrated redshift errors in order for systematic errors not to contribute to the dark energy error budget (though specific tolerances depend upon several details). This will be a relatively challenging goal for a photometric redshift calibration program to attain. DES, JDEM, EUCLID and LSST will all require calibration of very faint galaxies, where precise photo-z's are difficult to obtain. Moreover, the types of galaxies imaged, and for which spectra may be available, varies as a function of redshift, so some understanding of the details of galaxy evolution will be needed in order to achieve calibration goals.

It is natural to explore simple methods to sacrifice some of the statistical power of imaging surveys in order to mitigate larger systematic errors. One of the simplest techniques we can employ to limit the effect of catastrophic outliers is to place cuts on the range of photometric redshifts utilized to infer cosmological parameters (Bernstein and Huterer 13 have explored such cuts for a particular model of photo-z outliers).[6] The most damaging catastrophic errors are those that take galaxies near the median redshift of the survey and scatter them to significantly lower or higher redshifts, so it is sensible to explore the losses in statistical power incurred by excising galaxies at the low- and high-redshift ends of surveys.

We demonstrate the utility of photometric redshift excision in this section by exploring a class of simple excision algorithms. In particular, we cut out all galaxies with *photometric* redshifts greater than some value, $z^{\rm cut}_{\rm max}$, and smaller than some value $z^{\rm cut}_{\rm min}$. Figure 14 shows the statistical errors on w_0 and $w_{\rm a}$ as a function of $z^{\rm cut}_{\rm max}$ and $z^{\rm cut}_{\rm min}$ for our Wide survey, whose characteristics are similar to those expected from an LSST- or Euclid-like survey. The relative costs depend mildly upon survey parameters.

Excising galaxies with photometric redshifts lower than $z^{\rm ph} \sim 0.3$ results in only a $\sim 7\%$

[6] [137] also study the ability to employ photometric redshift cuts to mitigate the effects of catastrophic outliers, which became available on the Arxiv while we were submitting the paper associated with this chapter for publication.

increase in the statistical errors on dark energy parameters. Likewise, excising galaxies with $z^{\rm ph} \gtrsim 2.4$ results in only a $\sim 10\%$ degradation in w_0 and $w_{\rm a}$ constraints. Excising both of these regions of photometric redshift leads to a reduction in constraining power of $\lesssim 20\%$. Fig. 14 is a valuable itemization of the statistical losses incurred by redshift cuts and indicates that excising low- and high-redshift portions of the imaging surveys may be an effective method to mitigate the influence of catastrophic photometric redshift errors at little cost in statistical error.

While Fig. 14 quantifies the cost of excising regions of photometric redshift, the parametric complexity of catastrophic photo-z errors makes specific statements about the benefit of such cuts more difficult. In the case of a *localized* catastrophe that places galaxies erroneously in the excised high- or low-redshift ends of the survey, the induced bias can be nearly completely removed at the cost of the statistical degradation in Fig. 14. We have begun a preliminary study of the benefits of redshift excision, including the case of *uniform* catastrophes. In the case of our WIDE survey, excision can considerably reduce systematic errors induced even by the *uniform* catastrophe when the core is not well-calibrate ($N_{\rm spec} \lesssim 10^5$), but this strategy is only of marginal value in the limit of a well-calibrated core. We limit the present discussion to the itemization in Fig. 14 and relegate further study of redshift cuts and possible self-calibration of specific types of catastrophic error to a follow-up study.

3.4 CONCLUSIONS AND DISCUSSION

We have studied the potential systematic errors that may be induced in dark energy parameters inferred from forthcoming weak lensing surveys as a result of a population of source galaxies with photometric redshifts that deviate significantly from their true redshifts. We used a particular operational definition of catastrophic photo-z errors that is subtly distinct from the use of this term in some of the existing literature. Throughout this work, the term *catastrophic photometric redshift error* refers to cases in which photo-z estimates differ significantly from true redshifts, the nature of the error has not been identified or calibrated with an accompanying spectroscopic data set, *and* the outlier population has not been removed

reliably from the imaging data prior to the construction of shear correlation statistics. One way to interpret our results is as requirements for spectroscopic calibration of outliers and the completeness with which outlier galaxies must be culled from the data set in order to render systematic errors in dark energy parameters small.

In order to provide relatively general guidelines on the fidelity with which outlier photo-z's must be understood, we have taken an agnostic position on the nature of what types of catastrophic photometric redshift outliers may be realized in forthcoming imaging data. This eliminates the need to anticipate what types of photo-z errors may occur at very small fractional rates in order to assess their general influence on dark energy parameters. To be sure, there are reasonable guesses that can be made regarding the nature of photometric redshift errors and many algorithms exist that estimate redshifts from photometric data and refine estimates based upon comparisons with large, spectroscopic data sets [e.g., 17; 19; 30; 38; 39; 54; 83; 122; 139]. However, we have not adopted any particular template for photometric redshift outliers. Instead, we have studied two extreme limiting cases of catastrophic photometric redshift error.

In the first class of photometric redshift error, which we dubbed the *uniform* catastrophe, photometric redshifts are poorly constrained and scattered over a broad range [see, e.g., 38; 83, for examples of such features]. Photo-z errors resembling our uniform type must be well controlled. If such errors occur even for a relatively small fraction of galaxies near the median redshift of a given survey, the systematic errors induced on dark energy parameter estimators will be significant. Roughly speaking, we find that the error rate per galaxy must be maintained at $F_{\rm cat} \lesssim$ a few $\times 10^{-4}$. However, the uniform catastrophic error is a relatively simple variety so that self-calibration may well be feasible. One could resign oneself to the fact that such an error will occur and add the error rate $F_{\rm cat}$ (and perhaps other parameters such as $\Delta z_{\rm cat}$) to the set of nuisance parameters to be marginalized over. This self-calibration could eliminate the systematic error, but will broaden statistical errors. In future work we will explore possibility of self-calibrating particular catastrophic photo-z errors.

The second class of errors, which we refer to as *localized* catastrophes, takes source galaxies with particular true redshifts and assigns them photometric redshifts with a large bias but small scatter. Localized catastrophes have a broader range of possibilities and are

more difficult to deal with. Fig. 11, Fig. 13, and Eq. (12) constitute a blueprint for estimating the severity of a broad range of possible localized photometric redshift catastrophes. Quite generally, we find that the systematic errors they induce are sensitive to the scheme used to bin the source galaxies in photometric redshift. This suggests that an iterative scheme of re-binning may be an effective strategy for identifying and mitigating the influence of localized catastrophic photo-z errors.

In § 3.3.4 we studied a simple strategy to limit the systematics induced by catastrophic photo-z outliers. First, we showed that the statistical leverage of the highest redshift ($z \gtrsim 2.4$) and lowest redshift ($z \lesssim 0.3$) source galaxies on dark energy constraints is minimal. Eliminating all such galaxies from consideration in inferring dark energy parameters results in only a small increase in the statistical errors of dark energy equation of state constraints, but may eliminate some of the most severe systematic errors induced by localized catastrophic photo-z outliers. This implies that well-designed cuts on $z^{\rm ph}$ will likely be a powerful and general means to mitigate systematics associated with photo-z determination at a relatively small cost in statistical error.

The published work that is most closely related to the present work is Bernstein and Huterer [13]. Our work is an extension and generalization of their study. Overall, we reach the same broad conclusions where the two studies are commensurable. In particular, we find that catastrophic errors of the localized variety must be controlled such that the rate of errors per galaxy is $F_{\rm cat} \lesssim 10^{-3}$ if they are to induce tolerable systematic errors on dark energy parameters.

Our work differs from and complements Bernstein and Huterer [13] in several important ways. First, we have relaxed the assumption that the true redshift distribution of the outlier population perfectly traces that of the core population within individual Source photometric redshift bins (see Eq. 3.4). Our treatment of photometric redshift errors is independent of the photometric redshift binning (as such errors would be in practice), while the approach of Bernstein and Huterer [13] is limited to cases in which photometric errors both trace the galaxy distributions within the Source Bin and span the redshift range of the Source Bin. While contamination of the Target redshift bin is typically the larger source of induced systematic error, our generalization illustrates that the effects of modifications to the Source

Bin are non-negligible and in some cases these offsets contribute significantly to the systematic errors on dark energy parameters. Second, we have studied catastrophic errors in cases where the *core* photometric redshift distribution is not perfectly calibrated. Accounting for uncertainty in the core distribution turns out to be quite important: for a fixed catastrophe the magnitude of the induced systematic errors can vary by several orders of magnitude over a reasonable range of priors on the core distribution. Third, we have explored cases of correlated shifts in photo-z errors that span multiple tomographic redshift bins (which will occur in practice), the extreme example being the *uniform* error.

We conclude our discussion section by referring to interesting, tangential results given in the appendix. In the appendix, we discuss the effect of different models of the nonlinear evolution of cosmological density perturbations on photometric redshift calibration requirements. Weak lensing measurements take significant advantage of measurements on nonlinear scales in order to constrain cosmology. Previous work has utilized the Peacock and Dodds [141] formula [e.g. 118; 119]; however, we find that using the more recent and more accurate fit of Smith et al. [171] significantly reduces the need for independent calibration of photometric redshifts. We have used the Smith et al. [171] formula in the main body of this chapter. We refer the reader to the Appendices for further details.

3.5 SUMMARY

We have adopted a simple, agnostic approach to estimate the levels at which uncalibrated photometric redshift outliers must be controlled to maximize the dark energy constraints from the weak lensing components of forthcoming imaging surveys such as DES, LSST, EUCLID, and JDEM. We present results for three fiducial imaging surveys: a relatively near-term DES-like survey; a future survey with a high surface density of galaxies but a relatively small fractional sky coverage (DEEP); and a future survey with half-sky coverage and a lower galaxy surface density (WIDE). We considered two extreme cases of large, uncalibrated errors. In the case of a *uniform* photo-z catastrophe, we considered galaxies erroneously assigned photometric redshifts that are unrelated to their true redshifts. In the

case of a *localized* photo-z catastrophe, we considered the erroneous placement of a small fraction of galaxies in some range of true redshifts at significantly different *photometric* redshifts. To be specific, we assigned galaxies in some range of true redshifts of width Δz centered on a true redshift $z_{\rm cat}$ to photometric redshifts near $z_{\rm cat}^{\rm ph}$ that differ significantly from $z_{\rm cat}$. For each type of error and survey, we assessed the severity of the systematic errors on dark energy parameters that would be induced by catastrophic photometric redshift errors. Our primary results are as follows.

1. A photometric redshift error of the *uniform* variety that is relevant for galaxies near the median redshift of the imaging survey, must be limited to a fraction of galaxies $F_{\rm cat} \lesssim 5 \times 10^{-4}$ for DES or DEEP and $F_{\rm cat} \lesssim 2 \times 10^{-4}$ for WIDE, in order to induce systematic errors that are small compared to the statistical errors on w_0 and $w_{\rm a}$.
2. *Localized* catastrophic errors are most severe when they take some fraction of galaxies with true redshifts near the median survey redshift and assign them significantly higher or lower photo-z's. For DES, assignments to higher photo-z's are more severe than assignments to lower photo-z's while the opposite is true for WIDE and DEEP. However, the systematic errors induced by these two extremes differ by less than a factor of two in all cases.
3. Limiting the fraction of galaxies exhibiting *localized* catastrophes at all redshifts to $F_{\rm cat} \lesssim 3 \times 10^{-3}$ for DES or $F_{\rm cat} \lesssim 10^{-3}$ for WIDE or DEEP will render them unimportant. For localized catastrophes that occur over a range of true redshifts of width $\Delta z_{\rm cat}$ near the median survey redshift, the fractional error rate must be controlled such that $F_{\rm cat}(\Delta z_{\rm cat}/0.1) \lesssim 1 - 3 \times 10^{-3}$.
4. Imperfect knowledge of the photo-z distribution for the *core* sample of galaxies loosens these requirements for uncalibrated catastrophic outlier control as depicted in Fig. 13. Roughly speaking, core calibration with spectroscopic samples smaller than the statistical equivalent of $N_{\rm spec} \lesssim 10^5$ leads to significantly reduced catastrophic error control requirements. Of course, in practice catastrophic error control and core calibration will both improve as $N_{\rm spec}$ increases.
5. The statistical leverage of the highest redshift ($z \gtrsim 2.4$) and lowest redshift ($z \lesssim 0.3$) source galaxies on dark energy constraints is small. Eliminating all such galaxies from

consideration in inferring dark energy parameters results in a $\lesssim 20\%$ increase in the statistical errors on dark energy, but may eliminate the most severe systematic errors induced by localized catastrophic photo-z outliers.

6. In the appendix, we show that dark energy parameter forecasts that include photometric redshift uncertainty vary significantly depending upon the treatment of the nonlinearity in the matter power spectrum. In particular, using the Smith et al. [171] fitting form (as we do in the main text) leads to weaker photo-z calibration requirements than does the Peacock and Dodds [141] formula upon which the results of Ma et al. [119] are based. The Smith et al. [171] formula has been shown to be more accurate than Peacock and Dodds [141] suggesting that degradation due to photo-z uncertainty may be less than Ma et al. [119] forecast. Only a rigorous numerical study can determine this definitively.

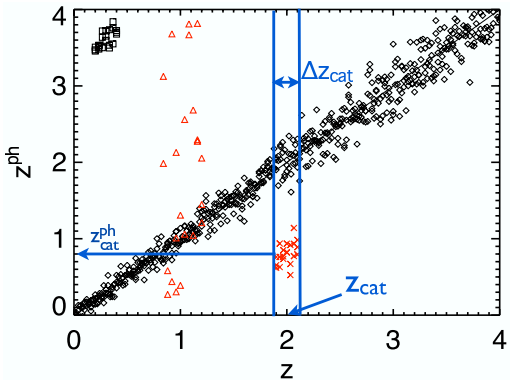

Figure 9: A toy illustration of a multi-component photometric redshift distribution. The aim of this figure is to provide a convenient, schematic representation of the photometric redshift distributions we explore. *Black diamonds* are galaxies in the primary peak of a Gaussian core population of photometric redshifts specified by Eq. (3.3). *Black squares* are galaxies in a secondary peak in a multi-modal core distribution. These photometric redshifts are offset from the line $z = z^{\rm ph}$, but they are a known component of the photometric redshift distribution, and if they are represented adequately in spectroscopic data they can be calibrated out. *Red crosses* are galaxies that reside in a catastrophic outlier population with significantly biased, but relatively localized, photometric redshifts. In our nomenclature, this population is not represented in spectroscopic calibration samples and contributes a systematic error to dark energy parameters. The *red triangles* represent galaxies that comprise a uniform catastrophic outlier population, where photometric redshifts are relatively unconstrained. The labels $z_{\rm cat}$, $\Delta z_{\rm cat}$, and $z^{\rm ph}_{\rm cat}$ designate the parameters of our catastrophic photometric redshift models.

Table 1: Representative Surveys and Baseline Constraints

Survey	$f_{\rm sky}$	N^A [arcmin^{-2}]	$z_{\rm med}$	$\sigma(w_0)$	$\sigma(w_{\rm a})$	$\sigma(w_{\rm p})$	$\mathcal{F}$
DES	0.12	15	0.7	0.25	0.77	0.07	18.6
WIDE	0.50	30	1.0	0.07	0.22	0.02	227.3
DEEP	0.05	100	1.0	0.10	0.33	0.04	75.6

NOTES.— Column (1) gives the survey that motivates the particular choice of parameters. Column (2) is the fractional sky coverage of the survey. Column (3) gives the effective galaxy number density N^A, in arcmin^{-2}. We have followed current convention and adopted a fixed shape noise of $\sqrt{\langle\gamma^2\rangle} = 0.2$, assuming deviations from this assumption to be encapsulated in the effective galaxy number density. Column (4) gives the median redshift of galaxies in the survey. Columns (5)-(8) give dark energy equation of state constraints in the limit of perfect knowledge of the photometric redshift distribution of sources. These include the uncertainty on the pivot equation of state $\sigma(w_{\rm p})$ and the product $\mathcal{F} = [\sigma(w_{\rm a}) \times \sigma(w_{\rm p})]^{-1}$. Note that these constraints are from the weak lensing components of these surveys only and account for statistical errors only.

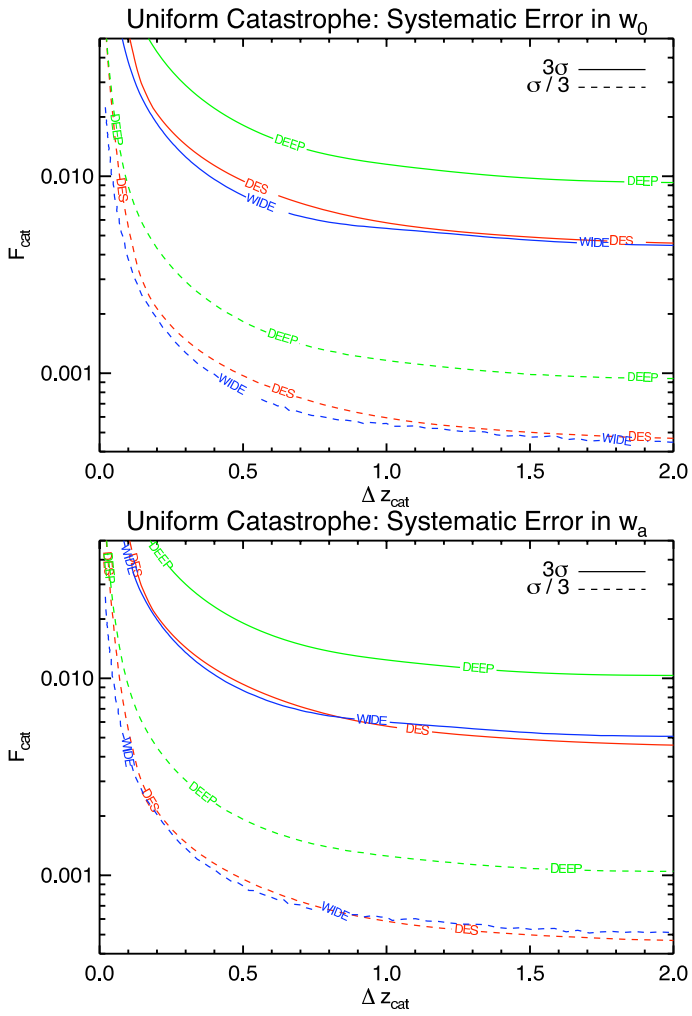

Figure 10: Systematic errors on dark energy parameters in the case of a *uniform* photometric redshift catastrophe. The horizontal axis is the width of the range in true redshift over which the uniform catastrophe is realized, $\Delta z_{\rm cat}$. This range in true redshiftis centered at $z_{\rm cat} = z_{\rm med}$ for each experiment ($z_{\rm med} = 0.7$ for DES and $z_{\rm med} = 1$ for DEEP and WIDE). The vertical axis is the catastrophic error rate per galaxy within this true redshift range, $F_{\rm cat}$. The *solid* (*dashed*) lines show contours of constant systematic error equal to three times (one third) the statistical errors on each of the dark energy parameters. The top panel shows contours for w_0 and the bottom for $w_{\rm a}$. Chance cancellations in the biases induced by high- and low-redshift galaxies cause the DES to be sensitive to catastrophic errors at similar levels to the WIDE survey and more sensitive than the DEEP survey.

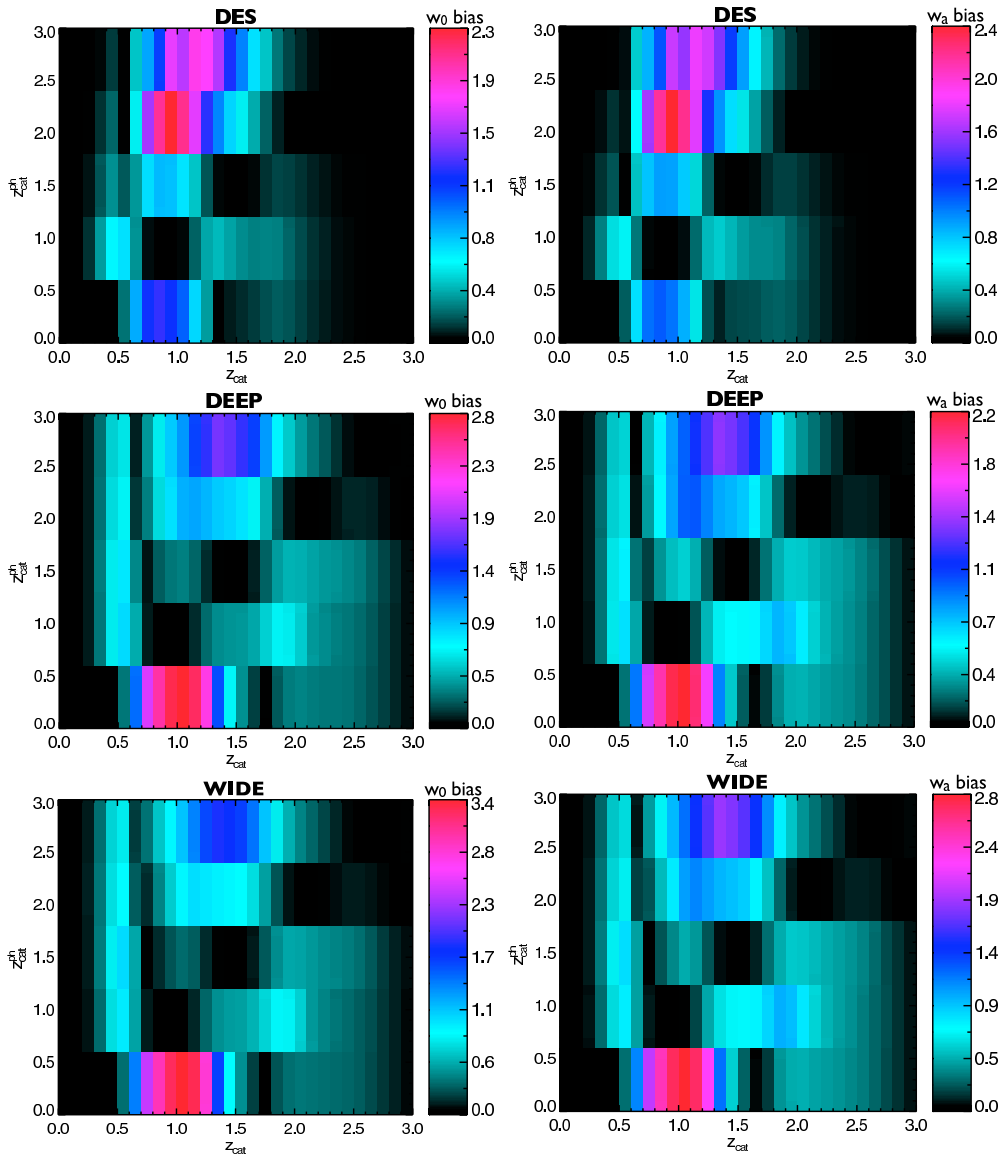

Figure 11: The severity of localized catastrophic errors as a function of the values of source $z_{\rm cat}$, and target $z_{\rm cat}^{\rm ph}$, of the catastrophic errors. Along the horizontal axes are the values of $z_{\rm cat}$ while the vertical axes show $z_{\rm cat}^{\rm ph}$, just as in Figure 9. Each point on this grid corresponds to a localized catastrophe with a fixed per-galaxy error rate of $F_{\rm cat} = 0.05$, and fixed values of both the photo-z spread $\sigma_{\rm cat} = 0.01$ as well as the width of the true redshift range over which the catastrophic error is made, $\Delta z_{\rm cat} = 0.05$. The effect of these catastrophes on w_0 is shown in the left column of panels, while the systematic error on $w_{\rm a}$ is shown in the right column. The absolute value of the induced systematic error is color coded; the numerical values labeling the color table to the right of each panel indicate the systematic error in units of the statistical uncertainty in the limit of perfect core calibration.

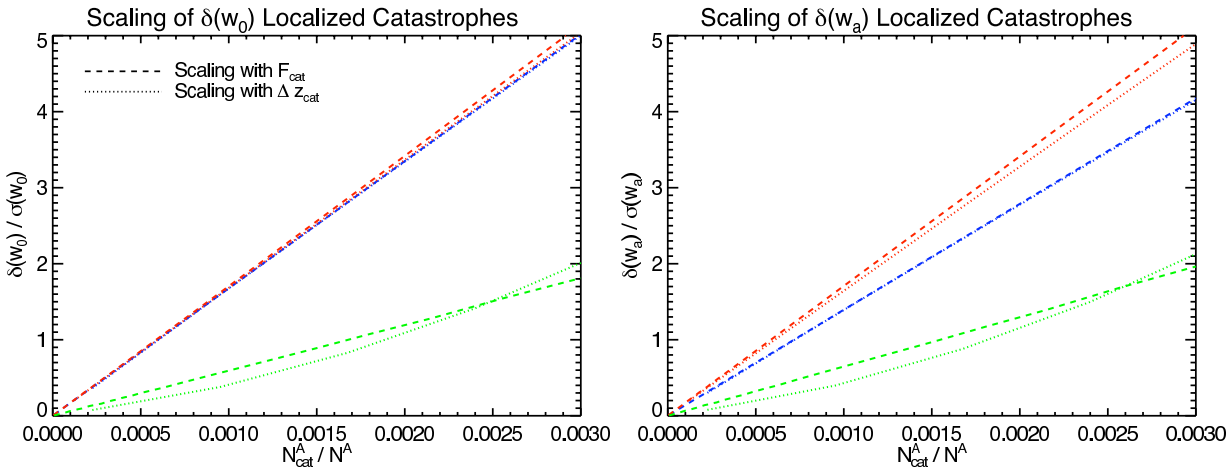

Figure 12: The scaling of systematic error in dark energy parameters with the fractional number density of sources whose photometric redshifts are catastrophically in error for our WIDE survey. On the vertical axes are the absolute value of the systematic error in w_0 (left panel) and $w_{\rm a}$ (right panel) in units of statistical uncertainty. On the horizontal axis is $N^{\rm A}_{\rm cat}/N^{\rm A} \approx F_{\rm cat} n(z_{\rm cat}) \Delta z_{\rm cat}$,where $n(z_{\rm cat})$ is the overall redshift distribution of sources. For the *dashed* curves, $\Delta z_{\rm cat} \equiv 0.1$ and we increase $N^{\rm A}_{\rm cat}$ by increasing $F_{\rm cat}$. These curves are all linear, as they should be. For the *dotted* curves $F_{\rm cat} \equiv 0.03$ and we increase $N^{\rm A}_{\rm cat}$ by increasing $\Delta z_{\rm cat}$. These curves grow approximately as the linear, *dashed* curves. Three different catastrophic error localizations are color coded as $(z_{\rm cat}, z^{\rm ph}_{\rm cat}) = (0.9, 0.3)$ in red, $(0.9, 2.7)$ in blue, and $(1.5, 2.7)$ in green. The difference in intrinsic severity between these outlier populations is reflected by the slope of the corresponding curves, with the steeper lines corresponding to the more severe systematic errors. The agreement between *dashed* and *dotted* lines for each outliers demonstrates that the systematic errors induced by sufficiently well-localized catastrophes ($\Delta z_{\rm cat} \lesssim 0.3$) scale approximately linearly with $\Delta z_{\rm cat}$ over an interesting range.

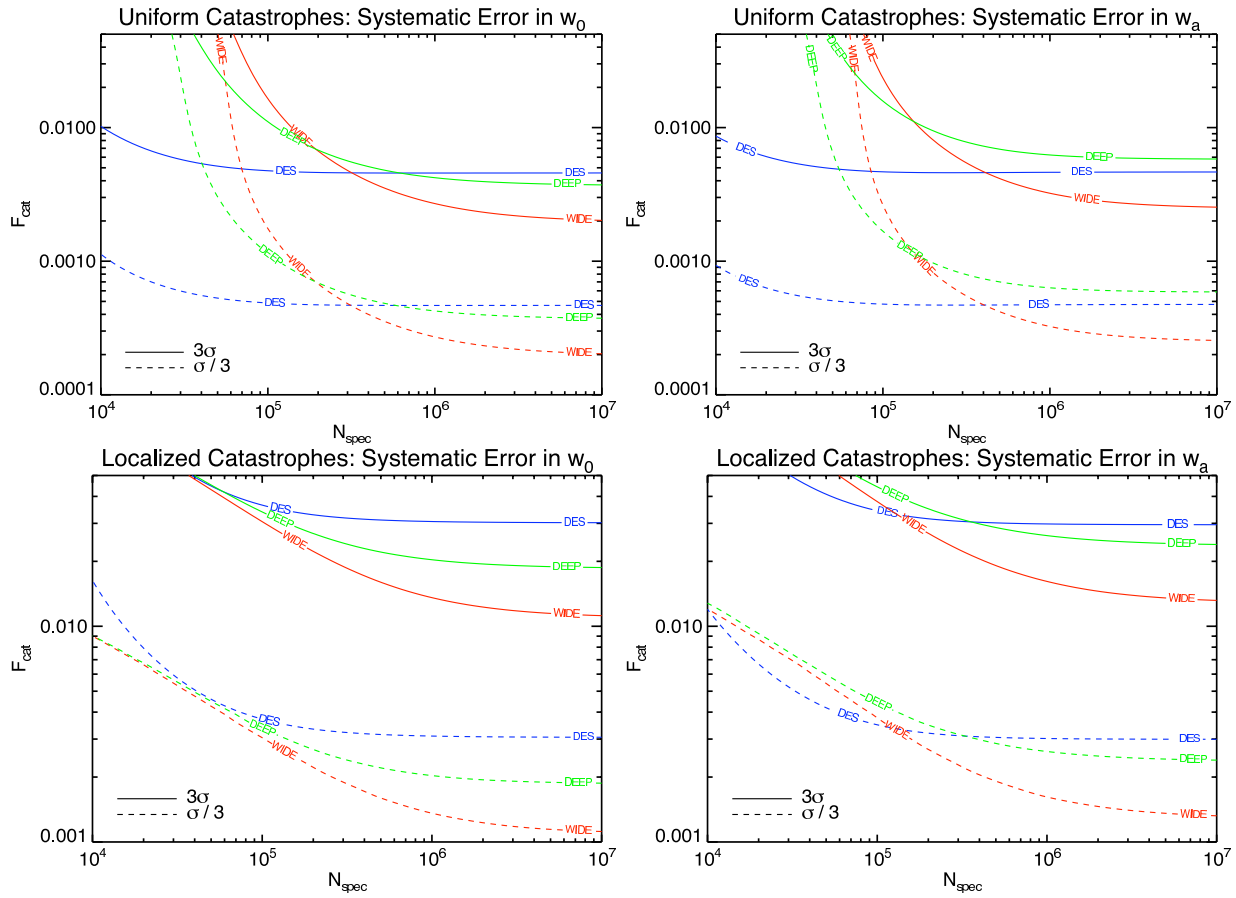

Figure 13: Contours of constant w_0 and w_a bias from the worst case catastrophe in units of the statistical uncertainty of the survey. Systematic errors in w_0 appear in the *left* panels and w_a in the *right* panels. Results for the worst case uniform catastrophe appear in the *upper* panels, and were generated with $\Delta z_{cat} = 1.5$ and $z_{cat} = 1.5$. Contours of systematic error produced by localized catastrophes appear in the *bottom* panels. Each of the localized contours have been calculated with $\Delta z_{cat} = 0.1$ and $\sigma_{cat} = 0.03$. For DEEP and WIDE $z_{cat} = 1.15$. and $z_{cat}^{ph} = 0.15$, and for DES $z_{cat} = 0.85$, and $z_{cat}^{ph} = 0.15$, in accordance with the results illustrated in Fig. 11.

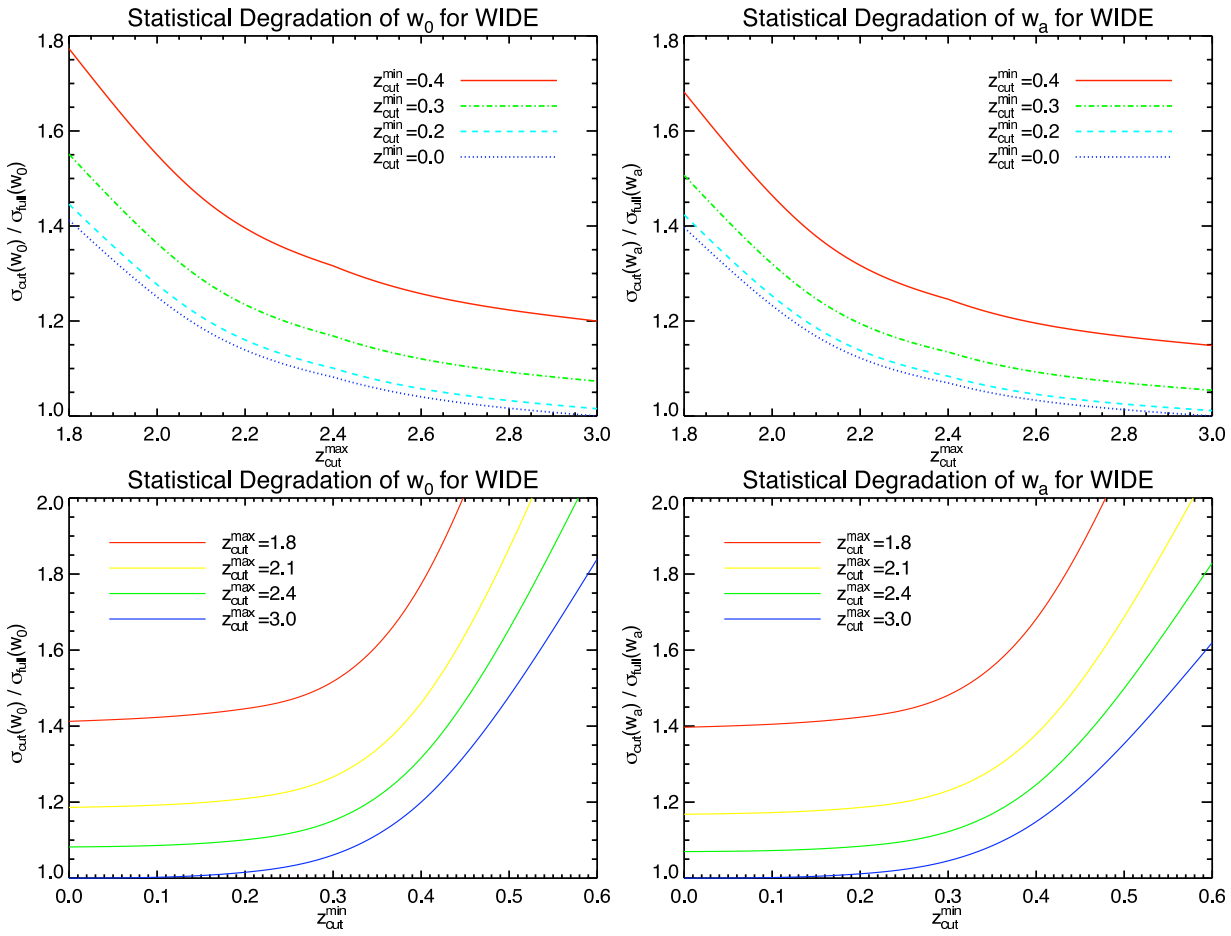

Figure 14: The statistical cost of excising low- and high-redshift shear information on constraints of w_0 (*left panels*) and w_a (*right panels*) for our WIDE survey. In the *top* row, the value of the maximum photometric redshift of the survey appears along the horizontal axis while the different lines show different choices of the minimum photometric redshift as indicated. Along the vertical axis is the fractional increase in dark energy parameter constraints relative to the constraints provided by a survey with our standard tomography. In the *bottom* row, the value of the minimum photometric redshift of the survey runs along the horizontal axis while the different lines show different choices of the maximum photometric redshift as indicated.

4.0 GENERAL REQUIREMENTS ON MATTER POWER SPECTRUM PREDICTIONS FOR COSMOLOGY WITH WEAK LENSING TOMOGRAPHY

4.1 INTRODUCTION

As demonstrated in previous chapters, weak gravitational lensing of galaxies by large-scale structure has the potential to provide a wealth of information about or cosmology [11; 46; 56; 71; 86; 87; 99; 143; 165; 187; 190]. Forthcoming imaging surveys such as the Dark Energy Survey (DES), the survey to be conducted by the Large Synoptic Survey Telescope (LSST), the survey of the European Space Agency's Euclid satellite [151], and the proposed Wide Field Infra-Red Survey Telescope (WFIRST) expect to exploit measurements of weak gravitational lensing of distant source galaxies to constrain the properties of the dark energy [4; 43; 64; 65; 70; 72; 76; 78; 84; 130; 148; 149; 173; 179; 180; 197; 198; 204]. Over the last several years, it has been recognized that the limited precision with which the matter power spectrum can be predicted may be one of several important, systematic errors that will need to be controlled in order to realize this goal (e.g., Refs. [25; 49; 58; 80; 89; 155; 166; 191; 197; 199]).

A large part of this uncertainty is due to the effects of baryons on lensing power spectra, which were largely neglected in much of the early literature on cosmological weak lensing. Several groups, including our own, have begun numerical simulation programs designed to address this issue with large-scale numerical simulations (e.g., Ref. [15; 26; 67; 68; 69; 166]). With the notable exception of Ref. [80], there have not been detailed studies of the precision with which the matter power spectrum must be predicted before it becomes a relatively small contributor to the error budget. In the run-up to large, computationally-intensive and

human resource-intensive simulation campaigns, we have studied the required precision with which the matter power spectrum must be predicted in order to realize anticipated dark energy constraints from cosmic weak lensing tomography. We present our results as a set of general guidelines on the systematic errors on the matter power spectrum as a function of scale and redshift.

Another potentially dominant source of error for dark energy parameter estimators arises from the necessity of using approximate redshifts determined from photometric data rather than spectroscopic redshifts [13; 82; 100; 108; 118; 119; 135; 178; 196; 203]. Photometric redshifts are significantly less precise than spectroscopic redshifts, and can exhibit large biases. Interestingly, we find that the precision with which the power spectrum must be predicted is very sensitive to the precision of photometric redshift determinations and vice versa, a result hinted at in the Appendix of Ref. [64]. As part of our analysis, we model photometric redshift uncertainty and show how the precision with which the power spectrum must be predicted varies with photometric redshift errors, and conversely.

Briefly, we find that if prior information on the photometric redshift distribution is weak, then dark energy constraints degrade $2-3$ times more rapidly with uncertainty in $\mathrm{P}_\delta(k)$ than if the photo-z distribution is characterized with high precision. Thus we find that when photo-z uncertainty is taken into account the calibration requirements on the theoretical prediction for the matter power spectrum are more stringent than previously thought. The complementarity of galaxy clustering statistics with weak lensing, well-studied in other contexts (see, for example, Refs. [198], [201], [168], [90], and [205]), has an ameliorating effect on power spectrum misestimations. Even when restricted to degree-scales, including galaxy correlation information can mitigate dark energy systematics induced by errors in the prediction for $\mathrm{P}_\delta(k)$ by up to 50%; alternatively, neglecting galaxy clustering statistics can cause the statistical constraints on dark energy parameters to degrade $2-5$ times more rapidly with uncertainty in either $\mathrm{P}_\delta(k)$ or the photo-z distribution.

This chapter is organized as follows. In §4.2 we describe how we model uncertainty in photometric redshifts and in $\mathrm{P}_\delta(k)$ as well as our methods for estimating statistical and systematic errors. We present our results in §4.3 and discuss their implications in §4.4. We conclude in §4.5 with a summary of our primary results.

4.2 METHODS

4.2.1 The Matter Power Spectrum

A significant amount of the constraining power of weak lensing surveys will come from scales on which the nonlinear effects of gravitational collapse cannot be neglected [81; 197; 200]. It is possible to excise data on relatively small scales, but such an excision significantly degrades dark energy constraints from cosmological weak lensing [81; 197]. Modeling nonlinear structure formation will be essential in order for forthcoming galaxy imaging surveys to realize their potential for constraining dark energy and modified gravity [63; 80]. As this modeling is uncertain and can be an important source of error, we study the errors induced on dark energy parameter estimators by uncertainty in the underlying matter power spectrum and we quantify the relative importance of theoretical power spectrum errors as a function of wavenumber. Our results may serve as a guideline for computational programs aimed at predicting accurate and precise matter power spectra for the purpose of comparing with weak lensing data.

In the current and past literature, the three most commonly-used techniques employed for predicting the matter power spectrum in the mildly nonlinear regime are the fitting formula of Peacock & Dodds [141], the Halo Model (see Ref. [33] for a review, as well as the many references therein), and the fitting formula of Smith et al. [171]. As a rough look at the contemporary level of uncertainty in predictions for the matter power spectrum $\mathrm{P}_\delta(k)$, and to set the stage for what is to come, we have plotted in Fig. 15 the fractional difference in $\mathrm{P}_\delta(k)$ at several different redshifts predicted by these three nonlinear evolution models. The differences in the predictions made by these methods become significant on scales ($k \gtrsim 0.2h\mathrm{Mpc}^{-1}$), which coincides with the scales at which the constraining power of weak lensing begins to peak. When the baryonic physics of galaxy formation is taken into account (for example, as in Refs. [58; 155; 166]), among other possible effects the matter distribution within halos is known to change relative to N-body (dark matter-only) simulations. To illustrate the effect such a rearrangement may have on the matter power spectrum, in Fig. 15 we additionally plot the fractional difference in $\mathrm{P}_\delta(k)$ that is induced when the Halo Model

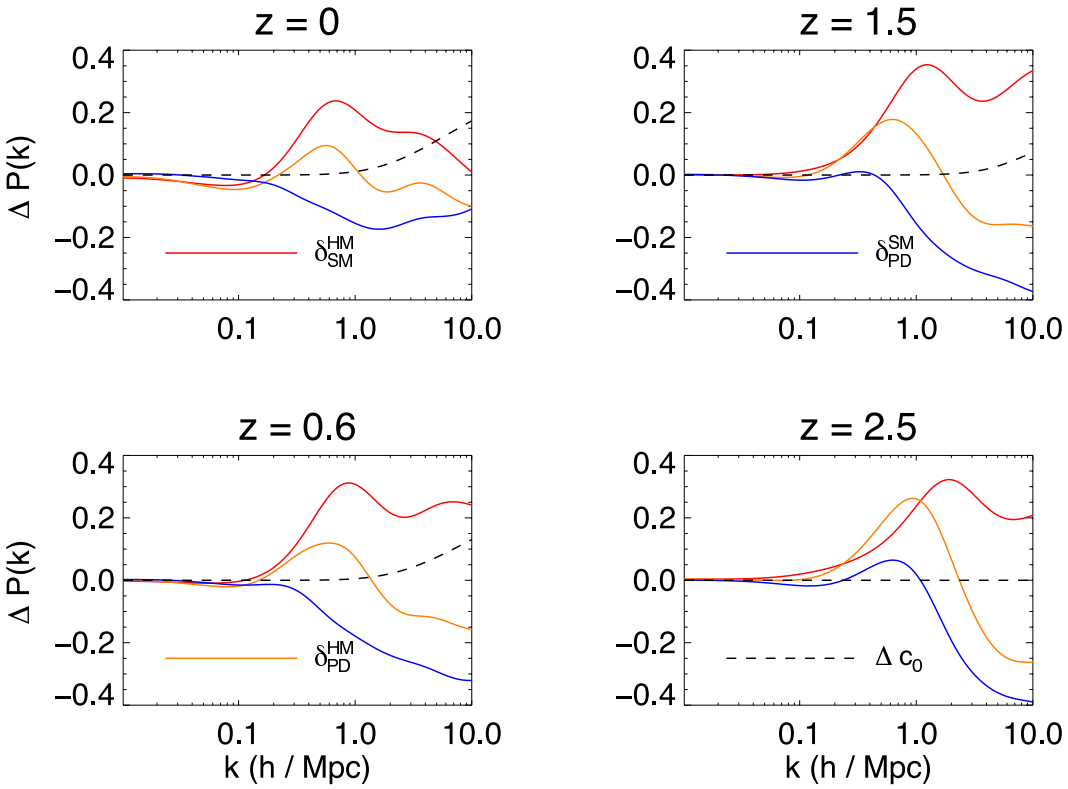

Figure 15: Fractional difference in $P_\delta(k)$ as predicted by the three different models of nonlinear evolution. The superscript indicates the fiducial model, the subscript the model inducing the fractional change. For example, the curve labelled by $\delta^{\rm HM}_{\rm SM}$ represents the fractional change to the $P_\delta(k)$ induced by using Smith et al. to compute nonlinear power rather than the fiducial halo model; thus when this curve is positive the Smith et al. prediction for $P_\delta(k)$ exceeds that of the halo model. The dashed curve labeled Δc_0 shows the fractional change to $P_\delta(k)$ induced by systematically misestimating the mean concentration of dark matter halos by 20%.

correctly predicts the nonlinear evolution but the mean concentration of dark matter halos (quantified by the parameter c_0 in Eq. 4.1) is misestimated by 20%.

In this chapter, we attempt to remain relatively agnostic about the types of errors that may be realized in predictions of the matter power spectrum at the percent and sub-percent levels, because such accurate predictions have not yet been made. We study two classes

of models for uncertainty in the matter power spectrum. Our first model is motivated by simulation results suggesting that one large source of systematic error in nonlinear power spectrum predictions may arise from a systematic error in predictions for the internal structures of dark matter halos. Significant rearrangement of dark matter may be a result of baryonic processes during galaxy formation, for example [58; 89; 155]. So, we suppose that halo abundances and halo clustering are well-known so that the halo model accurately predicts the gross shape of $\mathrm{P}_\delta(k)$, but that the distribution of matter within halos is uncertain. For our purposes, the average halo density profile may be taken as the Navarro-Frenk-White (NFW) density profile [133],

$$\rho(r) \propto \left(c\,\frac{r}{R_{200\mathrm{m}}}\right)^{-1} \left(1 + c\,\frac{r}{R_{200\mathrm{m}}}\right)^{-2}. \tag{4.1}$$

The parameter c describes the concentration of the mass distribution towards the halo center; the mass contained within the radius $R_{200\mathrm{m}}$ defines the halo mass. The average halo concentration varies with halo mass and evolves in redshift; we model these dependences as a power laws,

$$c(m, z) = c_0[m/m_{*,0}]^\alpha (1+z)^\beta, \tag{4.2}$$

where $m_{*,0} = 2.2 \times 10^{12}\ M_\odot/h$ is the typical mass of a halo that is just starting to collapse at $z = 0$. We allow the parameters α, β, c_0 to vary about their fiducial values of $c_0 = 10.7$, $\alpha = -0.11$, and $\beta = -1$, so that in this model uncertainty in $\mathrm{P}_\delta(k)$ stems exclusively from uncertainty in the distribution of mass within gravitationally-collapsed, self-bound objects. Our fiducial values for α and β are chosen to match results from the Millennium Simulation [134], but our fiducial value of c_0 is roughly a factor of two larger than the mean concentration of Millennium Simulation redshift-zero halos. This larger value is motivated by the results from Ref. [155] showing that baryonic physics typically produces a significant enhancement to the concentrations of dark matter halos. We thus set our fiducial c_0 according to Ref. [155] rather than the Millennium results.

Our second class of model for uncertainty in the matter power spectrum $\mathrm{P}_\delta(k, z)$, is significantly more general. After choosing a technique for predicting a *fiducial* matter power spectrum on scales of interest, we parameterize uncertainty in $\mathrm{P}_\delta(k)$ as follows. The range of scales between $k = 0.01h\mathrm{Mpc}^{-1}$ and $k = 10h\mathrm{Mpc}^{-1}$ are binned evenly in $\log(k)$, and

uncertainty in the matter power spectrum on a scale k in the $i^{\rm th}$ bin is parameterized via an additional parameter $\delta\ln(P_{\rm i})$, such that

$$\mathrm{P}_\delta(k,z) \rightarrow \mathrm{P}_\delta(k,z)\,(1+\delta\ln(P_{\rm i}))\,.$$

The parameters $\delta\ln(P_{\rm i})$ are then allowed to freely vary about their fiducial values of zero. We discuss our choice for the number of $\delta\ln(P_{\rm i})$ parameters in §4.2.4. This scheme for studying $\mathrm{P}_\delta(k)$ uncertainty was first introduced in the context of weak lensing in [80].

4.2.2 Photo-z Uncertainty

As uncertainty in photometric redshifts is likely to be one of the chief contributions to the error budget in future lensing measurements of dark energy, forecasts of dark energy parameter constraints need to include photo-z uncertainty to realistically estimate the constraining power of future surveys. While it is the primary goal of this chapter to study the influence of uncertainty in the matter power spectrum, our previous work has shown that photo-z calibration requirements can depend sensitively on the fiducial model for the nonlinear evolution of $\mathrm{P}_\delta(k)$. In the Appendix of Ref. [64], we showed that as prior information on the photometric redshift distribution decreases, constraints on w_0 degrade ~ 5 times faster if the Peacock & Dodds fitting formula is used to model nonlinear power rather than the fitting formula of Smith et al. (halo model results are intermediary between these two). The primary reason for this difference stems from Smith et al. predicting more small-scale power relative to the other two models, a trend that is apparent in Fig. 15. These results demonstrate that there is a nontrivial interplay between the nonlinear evolution of $\mathrm{P}_\delta(k)$ and the redshift distribution of the sources used to measure the weak lensing signal. This suggests that uncertainty in photo-z's and $\mathrm{P}_\delta(k)$ need to be treated simultaneously in order to accurately predict the calibration requirements for the matter power spectrum.

We model the underlying redshift distribution as $n(z) \propto z^2 \exp -(z/z_0)$, where the normalization is fixed so that $\int_0^\infty n(z)\mathrm{d}z = N_{\rm A}$, the mean surface density of sources in the survey. For weak lensing studies of dark energy, the most ambitious planned experiments for the next ten years will be LSST and Euclid; these surveys correspondingly have the

most stringent calibration requirements, and so we find it useful to quantitatively phrase our results for future very-wide-area surveys such as these. Thus unless stated otherwise we choose $N_{\rm A} = 30\,{\rm gal/arcmin}^2$, and $z_0 = 0.34$, corresponding to a median redshift of unity. We emphasize here, though, that the qualitative trends in all of our results are unchanged by the particular details of the survey characteristics.

We treat photometric redshift uncertainty in a relatively general manner, following the previous work of Ref. [119]. We assume that the source galaxies are binned according to *photometric* redshift and that the true redshift distributions of the galaxies within each of the photometric redshift bins, $n_i(z)$, are related to the overall, true galaxy source redshift distribution, $n(z)$, via

$$n_i(z) = n(z) \int_{z^i_{\rm low}}^{z^i_{\rm high}} {\rm d}z^{\rm ph} P(z^{ph}|z^{sp}),$$

where $z^i_{\rm low}$ and $z^i_{\rm high}$ are the boundaries of the $i^{\rm th}$ photometric redshift bin. Photo-z uncertainty is controlled by the function $P(z^{ph}|z^{sp})$. We take this to be a Gaussian at each redshift,

$$P(z^{ph}|z^{sp}) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{(z - z^{\rm ph} - z^{\rm bias})^2}{2\sigma_z^2}\right]. \tag{4.3}$$

This may seem to be overly restrictive, but as the mean, $z^{\rm bias}$, and dispersion, σ_z, of this distribution can vary with redshift themselves, this parameterization allows for a wide variety of possible forms for the functions $n_i(z)$. We adopt a fiducial model for photometric redshift error in which $\sigma_z = 0.05(1+z)$, while $z^{\rm bias} = 0$ at all redshifts. We use these fiducial functions to set the values of σ_z and $z^{\rm bias}$ at 31 control points, tabulated at intervals of $\Delta z = 0.1$ between z=0 and z=3. The values of $z^{\rm bias}$ and σ_z at each of these redshifts are free parameters in our forecasts, so that we model photo-z uncertainty with $2\times 31 = 62$ free parameters. This choice of binning allows for maximal degradation in dark energy constraint in the absence of prior information about the photometric redshift distribution of source galaxies (priors that would result from, for example, a photo-z calibration program). This indicates that our parameterization does not enforce correlations that yield better-than-expected dark energy constraints. The dark energy constraints in the absence of prior information reduce to the same constraints that would be obtained with no binning in photometric redshift.

To model uncertainty in the fiducial photo-z distribution, we introduce priors on the

values of the dispersion and bias at the $i^{\rm th}$ redshift control points, σ_i^z and $z_i^{\rm bias}$, respectively. These priors are

$$\Delta\sigma_i^z = \sigma_i^z \sqrt{\frac{1}{2N_{\rm spec}^{\rm i}}} \tag{4.4}$$

$$\Delta z_i^{\rm bias} = \frac{\sigma_i^z}{\sqrt{N_{\rm spec}^{\rm i}}}, \tag{4.5}$$

where $N_{\rm spec}^{\rm i}$ is the number of spectroscopic galaxies used in each of the 31 bins of width $\delta z = 0.1$ to calibrate the photo-z distribution. Our implementation of priors on the photo-z parameters is certainly simplistic. For example, we have further simplified our calculations by setting all of the $N_{\rm spec}^{\rm i}$ equal to each other, effectively assuming that the calibrating spectra are sampled equally in redshift, whereas in practice there will be looser constraints on sources at high redshift than at low redshift. However, because the details of how a realistic calibration program will proceed remains uncertain at the present time, we use this simple model for prior information and postpone a refinement of this parameterization until the exact set of spectra that will be used to calibrate LSST and Euclid is better known.

We emphasize here that $N_{\rm spec}$ provides a convenient way to specify a one-parameter family of photo-z priors through Eqs. 4.4 & 4.5. The quantity $N_{\rm spec}$ is likely not the true size of the spectroscopic calibration sample, but rather the equivalent size of a sample that fairly represents the color space distribution of the sources used in the lensing analysis.

4.2.3 Observables

We take the lensing power spectra of source galaxies binned by photometric redshift as well as the galaxy power spectra and galaxy-lensing cross spectra as observables that may be extracted from large-scale photometric surveys. The Limber approximation relates the power spectrum $\mathcal{P}_{s_i s_j}(k, z)$ associated with the correlation function of a pair of three-dimensional scalar fields, $\rm s_i$ and $\rm s_j$, to its two-dimensional projected power spectrum, $P_{\rm x_i x_j}(\ell)$:

$$P_{\rm x_i x_j}(\ell) = \int dz \frac{\rm W_i(z) W_j(z)}{\rm D_A^2(z) H(z)} \mathcal{P}_{\rm s_i s_j}(k = \ell / D_{\rm A}(z), z). \tag{4.6}$$

Eq. 4.6 essentially describes how the two-dimensional scalar fields $\mathrm{x_i}$ are observed as projections of the three-dimensional scalar fields, $\mathrm{s_i}$. The angular diameter distance function is denoted by D_A, and $H(z)$ is the Hubble parameter.

The weight function $\mathrm{W}_i(z)$ specifies the projection of the 3D source fields onto the 2D projected fields:

$$\mathrm{x_i}(\hat{\mathbf{n}}) = \int dz \mathrm{W}_i(z) \mathrm{s_i}(D_\mathrm{A}\hat{\mathbf{n}}, z). \tag{4.7}$$

For galaxy fluctuations, the weight function is simply the redshift distribution of galaxies in the i^{th} tomographic bin, $n_i(z)$, times the Hubble rate:

$$\mathrm{W}_\mathrm{i}^g(z) = H(z) n_i(z).$$

The weight function associated with fluctuations in lensing convergence is given by

$$W_\mathrm{i}^\kappa(z) = \frac{3}{2} H_0^2 (1+z) \Omega_\mathrm{m} \mathrm{D_A}(z) \int_z^\infty dz' \frac{D_\mathrm{A}(z,z')}{D_\mathrm{A}(z')} n_\mathrm{i}(z'),$$

where $D_\mathrm{A}(z, z')$ is the angular diameter distance between z and z'.

In principle, neither the redshift distribution of the galaxies used for the galaxy clustering nor the tomographic binning scheme need be the same as that used for cosmic shear sources, but for simplicity we use the same underlying distribution and binning for both so that the chief difference between the galaxy power spectrum $P_{g_\mathrm{i} g_\mathrm{j}}$, the convergence power spectrum $P_{\kappa_\mathrm{i}\kappa_\mathrm{j}}$, and the cross-spectrum $P_{\kappa_\mathrm{i} g_\mathrm{j}}$, is the form of the weight functions. Above and throughout, lower-case Latin indices label the tomographic redshift bin of the sources. For a survey with its galaxies divided into N_g redshift bins used to measure the galaxy clustering, and N_s bins for the galaxies used to measure cosmic shear, there will be $N_g(N_g+1)/2$ distinct 2-D galaxy power spectra $P_{g_\mathrm{i} g_\mathrm{j}}$, $N_s(N_s+1)/2$ distinct convergence power spectra $P_{\kappa_\mathrm{i}\kappa_\mathrm{j}}$, and $N_s N_g$ distinct cross-spectra $P_{\kappa_\mathrm{i} g_\mathrm{j}}$.

The matter power spectrum, $\mathcal{P}_{\mathrm{s_i s_j}}(k,z) = \mathcal{P}_\delta(k,z)$, sources the three-dimensional power in cosmic shear, whereas the source power for galaxy-galaxy correlations is the 3-D galaxy power spectrum $\mathcal{P}_{\mathrm{s_i s_j}}(k,z) = \mathcal{P}_{g_\mathrm{i} g_\mathrm{j}}(k,z)$. In all of our calculations we restrict galaxy correlation information to low multipoles $\ell \leq 300$; at redshift z=1 this corresponds to fluctuations of wavenumber $k \approx 0.2h\mathrm{Mpc}^{-1}$, so it will suffice for our purposes to use a linear, deterministic bias to relate the mass overdensity, $\delta(z)$, to the galaxy overdensity, $\delta_g(z) = b(z)\delta(z)$.

We allow for a very general redshift-dependent bias. To model uncertainty in the galaxy bias function $b(z)$, we allow the bias to vary freely about its fiducial value of unity in N_b galaxy bias bins, evenly spaced in true redshift, so that uncertainty in galaxy bias is encoded by N_b parameters. We computed dark energy constraints using the Fisher analysis technique described in §4.2.4 for N_b ranging from 1 to 30. We find that the dark energy constraints are insensitive to N_b ranging from $1 - 15$. Throughout this chapter, we present results pertaining to $N_b = 10$ bins, so that the value of the galaxy bias function $b(z)$ has independent, parametric freedom in redshift bins of width $\delta z = 0.3$. While finer binning is possible, particularly if the number of tomographic galaxy bins N_g is increased, a further increase of parametric freedom is unnecessary as galaxy bias is not a rapidly varying function of redshift (see, for example, Ref. [28]).

As a further simplification, we set the *fiducial value* of the bias function to unity at all redshifts, $\mathrm{b_i}(z) = b(z) = 1$ at all z for all i. This choice of fiducial parameter values is conservative, because the galaxies observed as part of high redshift samples will likely be biased [29], exhibiting relatively stronger correlations than matter, so we underestimate signal-to-noise of galaxy clustering measurements in the fiducial case.

4.2.4 Parameter Forecasting

We estimate the constraints from upcoming photometric surveys using the formalism of the Fisher information matrix. Useful references for this formalism include [4; 93; 103; 163; 181]. The Fisher matrix is defined as

$$F_{\alpha\beta} = \sum_{\ell_{\rm min}}^{\ell_{\rm max}} (2\ell + 1) f_{\rm sky} \sum_{\rm A,B} \frac{\partial \mathcal{O}_{\rm A}}{\partial p_\alpha} \mathbf{C}^{-1}_{\rm AB} \frac{\partial \mathcal{O}_{\rm B}}{\partial p_\beta} + F^{\rm P}_{\alpha\beta}. \tag{4.8}$$

The parameters of the model are p_α and the $\mathcal{O}_{\rm A}$ are the observables described in § 4.2.3. Greek indices label model parameters while Latin, upper-case indices label distinct observables. We take $\ell_{\rm min} = 2$ for all observables. For the lensing spectra, we take $P_{\kappa\kappa}$ observables we set $\ell_{\rm max} = 3000$ so that the assumptions of weak lensing and Gaussian statistics remain relatively reliable [32; 44; 165; 185; 192]. For the galaxy clustering statistics, we eliminate small-scale information so that we do not need to model scale-dependent galaxy bias, which

is potentially complicated in itself. Therefore, we set $\ell_{\rm max} = 300$ for P_{gg} and $P_{\kappa g}$, corresponding roughly to angular scales of ~ 1 degree. We emphasize that this restriction is very conservative as it implies that our joint analysis does *not* employ the use of Baryon Acoustic Oscillation features in our galaxy power spectra. As we will see, the added benefit of a joint analysis stems primarily from the increased ability to self-calibrate parameterized uncertainty in $\mathrm{P}_\delta(k)$ and the distribution of photometric redshifts.

In Eq. 4.8, $\mathbf{C}^{-1}_{\rm AB}$ is the inverse of the covariance matrix; our treatment of the covariance matrix calculation, and its associated Fisher matrix, is very similar to that in Ref. [73], to which we refer the reader for additional details. Briefly, the covariance between a pair of power spectra $P_{\rm x_i x_j}$ and $P_{\rm x_m x_n}$ is given by

$$\mathbf{Cov}(P_{\rm x_i x_j}, P_{\rm x_m x_n}) = \tilde{P}_{\rm x_i x_m}\tilde{P}_{\rm x_j x_n} + \tilde{P}_{\rm x_i x_n}\tilde{P}_{\rm x_j x_m}, \tag{4.9}$$

where in the case of either galaxy power or convergence power the observed spectra $\tilde{P}_{\rm x_i x_j}$ have a contribution from both signal and shot noise,

$$\tilde{P}_{\rm x_i x_j}(\ell) = P_{\rm x_i x_j}(\ell) + \mathrm{N}_{\rm x_i x_j},$$

where $\mathrm{N}_{g_i g_j} = \delta_{\rm ij}\mathrm{N}^{\rm A}_{\rm i}$ is the shot noise term for galaxy spectra, with $\mathrm{N}^{\rm A}_{\rm i}$ denoting the surface density of sources, and $\mathrm{N}_{\kappa_i \kappa_j} = \delta_{\rm ij}\gamma^2_{\rm int}\mathrm{N}^{\rm A}_{\rm i}$ is the shot noise for convergence. We calculate the observed cross-spectra $\tilde{P}_{\kappa_i g_j}$ without a contribution from shot noise, so that $\tilde{P}_{\kappa_i g_j} = P_{\kappa_i g_j}$, because galaxies are lensed by mass separated from them by cosmological distances, so the cross-correlation of the noise terms should be small. We adopt a common convention of setting the intrinsic galaxy shape noise $\gamma_{\rm int} = 0.2$ and absorb differences in shape noise between different surveys into the surface density of sources $N_{\rm A}$.

The inverse of the Fisher matrix is an estimate of the parameter covariance near the maximum of the likelihood, i.e. at the fiducial values of the parameters. Therefore, the measurement error on parameter α marginalized over all other parameters is

$$\sigma(p_\alpha) = \sqrt{[F^{-1}]_{\alpha\alpha}}. \tag{4.10}$$

Gaussian priors on the parameters are incorporated into the Fisher analysis via $F^{\rm P}_{\alpha\beta}$ in Eq. 4.8. If one is instead interested in unmarginalized errors, for example to test the intrinsic

sensitivity of the observables to a parameter p_α, the quantity $\sqrt{1/[F]_{\alpha\alpha}}$ provides an estimate of the uncertainty on p_α in the limit of zero covariance between p_α and any of the other parameters in the analysis.

The Fisher formalism can also be used to estimate the magnitude of a bias that would occur in parameter inference due to a systematic error in the observables. If $\Delta\mathcal{O}_{\rm A}$ denotes the difference between the fiducial observables and the observables perturbed by the presence of the systematic error, then the systematic offset in the inferred value of the parameters caused by the error can be estimated as

$$\delta p_\alpha = \sum_\beta [F^{-1}]_{\alpha\beta} \sum_\ell (2\ell+1) f_{\rm sky} \sum_{\rm A,B} \Delta\mathcal{O}_{\rm A} \mathbf{C}^{-1}_{\rm AB} \frac{\partial \mathcal{O}_{\rm B}}{\partial p_\beta}. \tag{4.11}$$

We assume a standard, flat ΛCDM cosmological model and vary seven cosmological parameters with fiducial values are as follows: $\Omega_{\rm m}h^2 = 0.13$, $w_0 = -1$, $w_{\rm a} = 0$, $\Omega_{\rm b}h^2 = 0.0223$, $n_{\rm s} = 0.96$, $\ln(\Delta^2_{\mathcal{R}}) = -19.953$, and $\Omega_\Lambda = 0.73$. We utilize the following marginalized priors: $\Delta\Omega_{\rm m}h^2 = 0.007$, $\Delta\Omega_{\rm b}h^2 = 0.001$, $\Delta n_{\rm s} = 0.04$, $\Delta\ln(\Delta^2_{\mathcal{R}}) = 0.1$. These priors are comparable to contemporary uncertainty [102], so this choice should be conservative. We have verified that strengthening these priors to levels of uncertainty that will be provided by Planck [77] does not induce a significant change to any of our results.

We determined the number of independent parameters for the matter power spectrum with an analysis of the off-diagonal elements of the inverse Fisher matrix. The parameter covariance is

$$\mathcal{Q}_{\alpha\beta} \equiv \frac{\mathcal{F}^{-1}_{\alpha\beta}}{\sqrt{\mathcal{F}^{-1}_{\alpha\alpha}\mathcal{F}^{-1}_{\beta\beta}}}. \tag{4.12}$$

In general, $-1 \leq \mathcal{Q}_{\alpha\beta} \leq 1$, with $\mathcal{Q}_{\alpha\beta} = (-)1$ corresponding to the case where parameters p_α and p_β are perfectly (anti-)correlated and $\mathcal{Q}_{\alpha\beta} = 0$ corresponding to uncorrelated parameters. By increasing the number of matter power spectrum parameters until the Fisher matrix is no longer invertible in the absence of prior information on these parameters, we determined that a lensing-only analysis reaches a level of total information loss when ten $\delta\ln(P_{\rm i})$ parameters are used. By studying the behavior of the off-diagonal Fisher Matrix entries as the number of matter power spectrum parameters are increased, we find that this

state of information loss occurs after values of $\mathcal{Q}_{\alpha\beta} = \pm 0.8$ obtain between pairs of distinct $\delta\ln(P_{\rm i})$, in agreement with the method used in Ref. [80] to arrive at this conclusion (D. Huterer, private communication). Because galaxy correlation observables provide additional information with which to self-calibrate matter power spectrum parameters, including $P_{\rm gg}$ and $P_{\kappa{\rm g}}$ allows for slightly finer binning in wavenumber, but for the sake of facilitating a direct comparison between the different sets of observables we have limited our analysis to ten parameters $\delta\ln(P_{\rm i})$, irrespective of whether we consider a joint analysis or weak lensing alone.

There is additional freedom in the choice of the number of tomographic bins one uses to divide both the sources used to measure lensing as well as the sources used to measure galaxy correlations. As has been noted in previous studies, for example Ref. [119], dark energy information from lensing saturates at $N_s = 5$ tomographic bins; this saturation point is determined by using the Fisher matrix (in the limit of perfect prior knowledge of all nuisance parameters, in our case the photo-z parameters σ_z^i and $z^i_{\rm bias}$, and the $\delta\ln(P_{\rm i})$ parameters) to compute the statistical constraints $\sigma(w_0, w_{\rm a})$ and increasing the number of tomographic bins until the constraints cease to improve. We find that for the case of galaxy clustering this information saturation occurs at $N_g = 10$ tomographic bins, although we note that this saturation point depends on the maximum multipole used in the analysis. For example, Ref. [201] uses $\ell_{\rm max} = 2000$ for their galaxy clustering analysis and finds that additional information is available by increasing the number of bins to $N_g = 40$. Our smaller N_g saturation point is a consequence of our conservative choice for $\ell_{\rm max}$, and the lack of BAO information implied by this choice.

4.3 RESULTS

4.3.1 Power Spectrum Self-Calibration

We begin by presenting our calculation of the scale-dependence of the sensitivity of weak lensing (with and without galaxy correlations) to the matter power spectrum. Our second

model for $\mathrm{P}_\delta(k)$ uncertainty is well-suited to this investigation: the constraints on parameter $\delta\ln(P_\mathrm{i})$ provide an estimate of the statistical significance of the weak lensing signal produced by correlations in the matter distribution on scales $k \approx k_\mathrm{i}$. In the prevailing jargon, this calculation corresponds to *self calibration* of the matter power spectrum.

The constraints from this computation are plotted as a function of scale in Fig. 16. The magenta curves at the bottom of the Fig. 16 pertain to *unmarginalized* constraints on the $\delta\ln(P_\mathrm{i})$. In other words, covariant uncertainty in cosmology, photometric redshifts, and galaxy bias is *not* taken into account in the magenta curves. In plotting the red and blue curves we illustrate our results when this covariance is accounted for by marginalizing over all other parameters in the analysis. The red curves correspond to a calculation with $N_\mathrm{spec} = 8000$, or $\Delta\sigma_z/\sigma_z \approx 10^{-2}$. The blue curves pertain to a calculation with $N_\mathrm{spec} = 2 \times 10^7$, which is sufficiently large that further increases to N_spec do not improve constraints on any of the parameters in our analysis, so priors this tight effectively correspond to the case where the photo-z distribution is known perfectly. The solid curves include galaxy correlation observables (P_gg and $P_{\kappa\mathrm{g}}$) in addition to lensing observables and thus lie strictly below the dashed curves (lensing only). The step-like appearance of the curves reflects the coarse binning in wavenumber of our parameterization of $\mathrm{P}_\delta(k)$ uncertainty: a forthcoming very-wide-area, LSST- or Euclid-like survey is only able to constrain ~ 10 independent matter power spectrum parameters (see § 4.2.4). While each of the curves in Fig. 16 pertains to a calculation in which we used the Smith et al. [171] fitting formula as our fiducial $\mathrm{P}_\delta(k)$, the results using either the halo model or the Peacock & Dodds [141] fitting formula are nearly identical, so conclusions drawn from Fig. 16 are quite robust to detailed changes in the fiducial model for nonlinear collapse.

The minimum of the unmarginalized constraints in Fig. 16 at $k \sim 2h\mathrm{Mpc}^{-1}$ occurs on the scale at which weak lensing is most intrinsically sensitive to matter overdensities. This minimum occurs on a physical scale nearly an order of magnitude smaller than the minimum of the marginalized constraints. This observation is an extension of the previous work of Ref. [80] and is itself an important result as it demonstrates the need for precision in the prediction for $\mathrm{P}_\delta(k)$ over the full range of nonlinear scales $k \lesssim 5\ h\mathrm{Mpc}^{-1}$. Because this shift in scale-dependence occurs even for the case of perfect prior knowledge on the photo-z

parameters ($N_{\rm spec} = 2 \times 10^7$), then it is not the effect of photometric redshift uncertainty that drives this shift in scale, but rather degeneracy with cosmological parameters. An analysis of the off-diagonal Fisher matrix elements shows that covariance of the $\delta\ln(P_{\rm i})$ with dark energy parameters is chiefly responsible for this dramatic shift in the scale-dependence. To be specific, with $p_{\rm i} = \delta\ln(P_{\rm i})$ and $p_{\rm j} = w_0, w_{\rm a}$, the corresponding $\mathcal{Q}_{\rm ij}$ (Eq. 4.12) have the maximum magnitudes of any of the cosmological parameters in our parameter set and they attain their maxima at $k \approx 0.1h{\rm Mpc}^{-1}$.

4.3.2 Statistical constraints on Dark Energy

We proceed with results on the sensitivity of dark energy constraints to uncertainty in predictions of the nonlinear evolution of $\mathrm{P}_\delta(k)$, incorporating possible additional uncertainty from photometric redshift errors. Results for the $\delta\ln(P_{\rm i})$ model appear in 4.3.2.1 while those pertaining to the more restrictive 'halo model" treatment of power spectrum uncertainty are discussed in 4.3.2.2.

4.3.2.1 The $\delta\ln(P_{\rm i})$ Model In Figure 17 we depict contours of the degradation in the statistical constraints on the dark energy equation of state associated with simultaneous uncertainty in $\mathrm{P}_\delta(k)$ and photometric redshifts. The constraints on w_0 and $w_{\rm a}$ are shown in units of the perfectly calibrated limit, when power spectra and photometric reshifts are known so well as to be inconsequential to the dark energy error budget; we denote the constraints on w_0 and $w_{\rm a}$ in the limit of perfect calibration by $\sigma^{\rm perf}(w_0)$ and $\sigma^{\rm perf}(w_{\rm a})$, respectively. For example, we plot the ratio $\Xi = \sigma(w_0)/\sigma^{\rm perf}(w_0)$ to illustrate the level of degradation of w_0 constraints. For the sake of scaling our results to an absolute statistical constraint, we summarize these baseline constraints in Table 2. We refer the reader to § 4.5 for a discussion of how the results we present here change when considering a survey with characteristics similar to DES.

The levels of constraint degradation depend upon the precision of both the matter power spectrum, parameterized by $\delta\ln(P_{\rm i})$, and the photometric redshift distributions of sources. We quantify the precision of power spectrum prediction by a prior constraint on the $\delta\ln(P_{\rm i})$

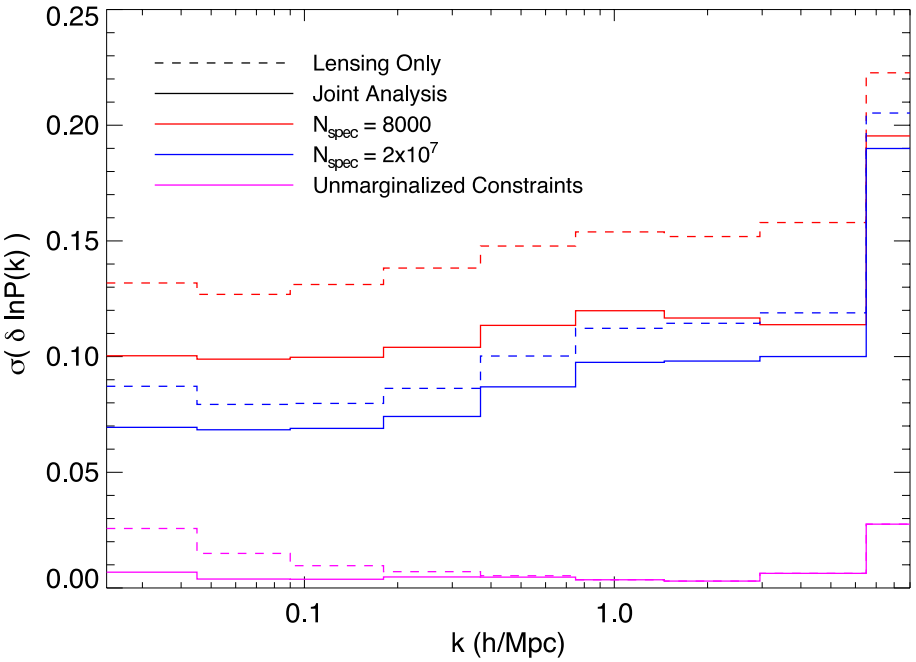

Figure 16: Statistical constraints on $\delta\ln(P_{\rm i})$, the parameters encoding uncertainty in the calibration of the matter power spectrum, are plotted against the scale of the wavenumber. We plot unmarginalized constraints as magenta curves; the minimum of the magenta curves at $k \approx 2\ h{\rm Mpc}^{-1}$ illustrates that the intrinsic sensitivity of the weak lensing signal peaks at this scale. With the red and blue curves we plot marginalized constraints on the $\delta\ln(P_{\rm i})$ parameters for different levels of uncertainty in the distribution of photometric redshifts. The vertical axis values for the red and blue curves give the statistical precision with which a future very-wide-area survey such as LSST or Euclid will be able to self-calibrate the theoretical prediction for the matter power spectrum on a given scale.

parameters, $\Delta\delta\ln(P_{\rm i})$. For example, a value of $\Delta\delta\ln(P_{\rm i}) = 0.1$ corresponds to a 10% precision on the bandpower at wavenumber $k_{\rm i}$. For simplicity, we apply the same prior at all wave bands in order to produce Fig. 17. The assumed level of calibration of photometric redshifts is specified by the parameter $N_{\rm spec}$, as discussed in § 4.2.2.

The horizontal gray lines roughly bound the range of precision in the prediction for $\mathrm{P}_\delta(k)$ on scales relevant to lensing and large-scale galaxy clustering that may be attainable

by near-future numerical simulation campaigns. The precise values of the realized precisions will depend upon the resources dedicated to address this issue as well as the ability of data to constrain baryonic processes that alter power spectra and numerical simulations to treat these baryonic processes. For example, the Coyote Universe simulation campaign [69] has already achieved a 1% calibration of the matter power spectrum to scales as small as $k \approx 1$ $h\mathrm{Mpc}^{-1}$; future results from, e.g., the Roadrunner Universe [60] will improve upon these results, although it is still not clear in detail how precisely $\mathrm{P}_\delta(k)$ will be calibrated to scales as small as $k \approx 10$ $h\mathrm{Mpc}^{-1}$, especially when uncertainty in baryonic physics is taken into account.

Fig. 17 contains contours for parameter degradation when using lensing data alone (dashed curves) as well as the corresponding constraint degradations when both lensing and galaxy clustering observables are considered (solid curves). To be sure, the additional information available to a joint analysis guarantees that $\sigma^{\mathrm{perf}}(w_0, w_\mathrm{a})$ from the joint galaxy clustering and lensing analysis is less than the corresponding constraint when considering lensing observables alone. In each case, we plot the degradation $\Xi = \sigma(w_0, w_\mathrm{a})/\sigma^{\mathrm{perf}}(w_0, w_\mathrm{a})$ relative to the idealized constraints for that technique. Accounting for this difference, Fig. 17 highlights the dramatic relaxation on the calibration requirements of both power spectrum predictions and photometric redshifts provided by including galaxy clustering statistics. For the purpose of studying dark energy, this relaxation is the most significant advantage provided by utilizing galaxy correlations in the analysis.

To further illustrate this point, consider particular examples that can be gleaned from Fig. 17. In the limit of perfect prior knowledge of $\mathrm{P}_\delta(k)$ and the photo-z distribution, the increase in constraining power that a joint analysis has over an analysis that includes lensing observables only is merely $\sim 15 - 20\%$; however, as the precision in the calibration of both photometric redshifts and the matter power spectrum decreases, the dark energy constraints degrade by a factor of $2 - 5$ more rapidly when galaxy clustering information is neglected. For example, suppose that predictions for the matter power spectrum attain 1% level of precision in advance of LSST or Euclid. A joint analysis with the statistical equivalent of $N_{\mathrm{spec}} \approx 20,000$ (a realistic proposition as we are assuming these calibrating galaxies to be distributed evenly in redshift) yields constraints on w_a that are weakened by 75% due

to photometric redshift uncertainty, whereas if galaxy clustering information is neglected $N_{\rm spec} \approx 100,000$ will be required to protect against the same level of degradation in dark energy constraints.

4.3.2.2 The Halo Model In Figure 18 we display the degradation in the statistical constraints on w_0 (left panels) and $w_{\rm a}$ (right panels) due to uncertainty in the halo concentration parameters c_0, α, and β [Eq. (4.2)]. The precision level of the calibration of c_0 appears as the horizontal axis in the top panels, α in the middle panels, and β the bottom panels. Along each row of panels in Fig. 18, only the parameter labeled on the horizontal axes has uncertain prior knowledge. In other words, the remaining two concentration parameters are treated as perfectly known for simplicity. As in Fig. 17, each curve is normalized to the constraint that would be realized in the limit of perfect knowledge of the halo structure parameters, $\sigma^{\rm cperf}$. Note however, that this baseline constraint has been recomputed for each assumed value of photometric redshift uncertainty, as specified by the color-coding of each curve, so that the degradation represents only that amount of additional degradation due to uncertainty about halo structure. Thus for each level of photometric calibration precision appearing in the legend, $N_{\rm spec} = 5000$, 10^4, 10^5, and 10^6, the vertical axis value gives the degradation in the $w_0(w_{\rm a})$ constraints strictly due to uncertainty in the parameter labeled on the horizontal axis.

In all cases, the degradation of dark energy parameters is relatively modest. In particular, the degradation induced by halo structure uncertainty alone is $\lesssim 15\%$ for all reasonable models, and significantly less if photometric redshifts are well calibrated or if galaxy clustering statistics are employed. This is in qualitative agreement with Ref. [197], who studied calibrating halo structure parameters as a means to account for the influence of baryonic processes on the lensing power spectrum. It is not expected that halo structure alone can account for all of the effects of baryonic processes on the power spectrum (e.g., Refs. [155; 166; 197; 199]). However, combining the results of Fig. 17 with Fig. 18 suggests that if uncertainty in the power spectrum due to baryonic processes can be modeled by concentrations with a residual of order $\sim 1\%$, then dark energy parameter degradation induced by such uncertainty may be limited to quite modest values. It may even be possible to incorporate additional parameters

to describe, for example, the hot gas components of groups and clusters [155; 166; 199], at a relatively modest statistical cost.

4.3.3 Systematic Errors on Dark Energy

In this section we explore the related problem of systematic errors on the dark energy equation of state parameters induced by uncertainty in predictions of the matter power spectrum. To this point, we have already addressed degradation in dark energy parameters induced by treating uncertainty in the matter power spectrum as a statistical uncertainty. In such a calculation, the underlying assumption is that a model for the power spectrum is accurate but the parameters of the model are known with imperfect precision. Our aim in this section is to elucidate dark energy equation of state errors in the related circumstance of a systematic error on power spectrum predictions. The underlying framework is that a model for the power spectrum exists and is assumed to be correct (or at least, that it contains the true power spectrum within its parameter set), but the model misestimates the power spectrum over some range of wavenumbers. This situation may be the most relevant to forthcoming data analyses if, as a specific example, the dominant errors in power spectrum predictions stem from systematic errors in the numerical treatment of baryons. The result of parameter inferences that use such systematically-offset theoretical power spectra will be systematically-offset dark energy equation of state estimators. In this section, we quantify the systematic errors in dark energy equation of state parameters induced by systematic offsets in the matter power spectrum. Not surprisingly, the results of the previous section will be a useful guide for anticipating and understanding the results of this section. As before, we begin with our more general $\delta\ln(P_{\rm i})$ model and later address systematic errors in the halo-based model.

4.3.3.1 The $\delta\ln(P_{\rm i})$ Model We begin our treatment of systematic errors with our $\delta\ln(P_{\rm i})$ model. This model treats the predicted power spectrum as a sequence of bandpowers distributed evenly in $\log(k)$. The generality of this model enables us to address a specific, important concern, namely: How does the severity of the induced systematic error in the

dark energy equation of state depend on the comoving scale k at which the prediction for $\mathrm{P}_\delta(k)$ is erroneous? In so doing, we can prescribe how well the theoretical matter power spectrum should be calibrated as a function of scale, setting a specific goal for large-scale simulation efforts.

In Fig. 19 we have plotted the systematic error induced on w_0 (top panel) and w_a (bottom panel) by a 5% error in $\mathrm{P}_\delta(k)$ at the comoving scale k labeled along the horizontal axis. The induced systematic errors have been calculated according to Eq. 4.11, where the systematic shift to the observables is induced by introducing a 5% systematic error in $\mathrm{P}_\delta(k)$ over a range of wavenumbers with a width that is 10% of k_err, the scale at which the error is centered, viz.

$$\frac{\Delta \mathrm{P}_\delta(k)}{\mathrm{P}_\delta(k)} = \begin{cases} 0.05 & : 0.95k_\mathrm{err} \leq k \leq 1.05k_\mathrm{err} \\ 0 & : \text{otherwise} \end{cases} \tag{4.13}$$

The curves in Fig. 19 are color-coded according to the priors on the photo-z distribution, and the magnitude of the induced systematic error has been normalized by σ^cperf, the statistical uncertainty of the parameter assuming perfect prior knowledge of each of the $\delta\ln(P_\mathrm{i})$ but with photo-z uncertainty at the level given by the color-coding of the curve and the observables included in the analysis.

Each curve approaches zero at large and small scales. Very large scale lensing correlations contribute little constraining power on dark energy, and so it should be expected that scales larger than $k \lesssim 10^{-2}\ h\mathrm{Mpc}^{-1}$ should contribute comparably little to the error budget; errors in $\mathrm{P}_\delta(k)$ on scales smaller than $k \gtrsim 5h\mathrm{Mpc}^{-1}$ are comparably tolerable because we make no use of correlation information from multipoles greater than $\ell_\mathrm{max} = 3000$. From the wavenumber-range containing the peaks in curves in Fig. 19 we can see that the most significant dark energy biases induced by systematic errors in $\mathrm{P}_\delta(k)$ come from errors on scales $0.05\ h\mathrm{Mpc}^{-1} \lesssim k \lesssim 3\ h\mathrm{Mpc}^{-1}$. This finding is consistent with the results presented in Fig. 16, in which we can see that this is the same range of scales over which the tightest statistical constraints can be obtained by self-calibrating the parameters $\delta\ln(P_\mathrm{i})$. Both of these figures thus illustrate that the weak lensing information about dark energy that will be available to future very-wide-area surveys such as LSST or Euclid comes primarily from gravitational lensing events produced by perturbations on scales $0.05\ h\mathrm{Mpc}^{-1} \lesssim k \lesssim 3$

$h\mathrm{Mpc}^{-1}$, in good agreement with previous results [80].

Although almost all of the curves are limited to systematic biases $\delta(w_0, w_a) \lesssim 1\sigma(w_0, w_a)$, we remind the reader that the y-axis value gives the magnitude of the bias when the systematic error in $\mathrm{P}_\delta(k)$ is isolated to just a single bin of wavenumbers of width $\Delta k/k = 0.1$, *and* when the magnitude of the power spectrum error is 5%. However, we only chose these particular values for the sake of making a definite illustration, and so for the results in Fig. 19 to be useful to the calibration program it will be necessary to scale the systematics we predict according to the particular details of the matter power spectrum error whose consequences are being estimated. We give several examples of this below to illustrate the utility of our calculations.

Suppose there is a 3% error in $\mathrm{P}_\delta(k)$ made over a range of $\Delta k/k = 0.4$, centered at $k \approx 2\ h\mathrm{Mpc}^{-1}$. For definiteness, consider a lensing-only analysis with photo-z uncertainty modeled by $N_{\mathrm{spec}} = 10^3$. The y-axis value of the corresponding (dashed, red) curve at $k \sim 2\ h\mathrm{Mpc}^{-1}$ is $\delta w_0 = 0.9\sigma(w_0)$. This value needs to be rescaled by a value of 3/5 to account for the difference between the magnitude of this example's actual error in $\mathrm{P}_\delta(k)$ and the 5% error plotted in Fig. 19; additionally, scaling by a factor of 4 is necessary to account for the fact that the range of scales over which the error is operative spans 4 of our bins[1], giving an estimate of $\delta w_0 \approx 2.2\sigma(w_0)$.

If a 3% $\mathrm{P}_\delta(k)$ error, again spanning 4 of our $\Delta k/k = 0.1$ bins, is instead made at $k \approx 0.8\ h\mathrm{Mpc}^{-1}$, the estimation method of the first example naively implies that there is zero systematic error associated with such a power spectrum misestimation because the error is centered at a wavenumber at which δw_0 changes sign, and so the contributions to the net dark energy systematic to the left and right of $k \approx 0.8\ h\mathrm{Mpc}^{-1}$ appear to cancel. Such cancellations are not necessarily spurious, and in fact a very general formalism for choosing a set of nuisance parameters specifically designed to take advantage of this phenomenon has recently been proposed [138]. However, because it may not be known if the sign of the matter power spectrum error inducing the biases also changes sign over the range of wavenumbers

[1]Simple, linear scaling is a very good approximation when correcting for the magnitude of the $\mathrm{P}_\delta(k)$ error, but this prescription is only approximately correct when rescaling according to the width of the range of scales over which the error is made. We find that the derivatives of lensing power spectra with respect to $\delta\ln(P_i)$ parameters are stable to roughly factor-of-five changes in numerical step-size, and so simple, linear scaling will be appropriate so long as the width of the wavenumber range is less than $\Delta k/k \lesssim 0.5$.

on which the error is made, it may not be possible to exploit this sign change to minimize the net effect of the error. In such a case, we advocate assuming the worst-case scenario, that biases produced by errors in multiple bins in wavenumber conspire to contribute additively; by construction this will yield a conservative estimate for the systematic induced by the error in $\mathrm{P}_\delta(k)$. Thus for this example, the absolute value of the y-axis values should be used to estimate the net dark energy systematic. The magnitude of the dashed, red curve for this particular systematic peaks at $\delta w_0 = 0.5\sigma(w_0)$ at the endpoints of its operative range, $0.6\ h\mathrm{Mpc}^{-1} \lesssim k \lesssim 1\ h\mathrm{Mpc}^{-1}$; the curve approaches these maxima from its value of zero at $k = 0.8\ h\mathrm{Mpc}^{-1}$ and so we approximate this as an $0.25\sigma(w_0)$ error spanning 4 of our bins. Thus for a 3% systematic error in $\mathrm{P}_\delta(k)$ spanning the range $0.6\ h\mathrm{Mpc}^{-1} \lesssim k \lesssim 1\ h\mathrm{Mpc}^{-1}$, our final estimate is given by $\delta w_0 = [0.25 \times 4 \times (3/5)]\,\sigma(w_0) = 0.6\sigma(w_0)$.

For power spectrum errors made over very broad ranges of wavenumber, the simple linear scaling of $\Delta k/k$ is no longer appropriate and one must rely on the full machinery of our calculation to integrate the absolute value of $\delta w_0/\sigma(w_0)$ over the scales over which the error is operative. This is also useful to conservatively estimate an ultimate target goal for the matter power spectrum calibration effort. When performing this integration on scales $1h\mathrm{Mpc}^{-1} \lesssim k \lesssim 5h\mathrm{Mpc}^{-1}$, the worst-case estimate is a systematic error of $\delta(w_0) \approx 3 - 4\sigma(w_0)$;[2] if one instead integrates a 5% $\mathrm{P}_\delta(k)$ error over the entire range of wavenumbers $0.01h\mathrm{Mpc}^{-1} \lesssim k \lesssim 5h\mathrm{Mpc}^{-1}$, a worst-case estimate of the coherently contributing biases ranges from $9 - 13\sigma$. To ensure that dark energy systematics are kept at or below the level of statistical constraints, 0.5% accuracy in the prediction for $\mathrm{P}_\delta(k)$ over the entire range of $0.01h\mathrm{Mpc}^{-1} \lesssim k \lesssim 5h\mathrm{Mpc}^{-1}$ will be required of the simulations calibrating the matter power spectrum. Note that these requirements are somewhat more restrictive than estimations from previous work [80].

4.3.3.2 Dependence on Multipole Range The results presented above in §4.3.3.1 depend sensitively on ℓ_{max}, the maximum multipole used in the cosmic shear analysis. Naturally, as ℓ_{max} increases the matter power spectrum must be modeled with greater precision and to smaller scales. Figure 20 represents a simple illustration of this point. Each curve in

[2]The exact number depends on the level of photo-z calibration as well as the choice of observables.

Fig. 20 pertains to a cosmic shear-only experiment in the limit of perfect knowledge of the photo-z distribution, but with different choices for $\ell_{\rm max}$ color-coded according to the legend. The axes are the same as those in the top panel of Fig. 19, and so this plot shows how w_0 biases induced by $\mathrm{P}_\delta(k)$ errors change with the choice for the maximum multipole used in the lensing analysis.

At the end of §4.3.3.1 we described how to use results such as those appearing in Fig. 20 to estimate the precision to which the matter power spectrum must be predicted in order to guarantee that dark energy biases induced by $\mathrm{P}_\delta(k)$ systematics are kept at or below the level of the statistical constraints. Briefly, one assumes the worst case scenario, that the sign of the $\mathrm{P}_\delta(k)$ errors conspire to contribute coherently to the dark energy bias. In this case, to estimate the most severe dark energy bias that could be induced by an error in $\mathrm{P}_\delta(k)$ made over a range of wavenumbers, one simply adds the absolute value of the relevant curve over the relevant range of wavenumbers. To obtain a more optimistic estimation, one could suppose that the errors in each bin are perfectly uncorrelated, in which case they may be treated as independent Gaussian random variables so that their net contribution to the error budget is computed by adding the individual contributions in quadrature. In either the pessimistic or optimistic case, one obtains a precision requirement by finding the magnitude of the $\mathrm{P}_\delta(k)$ error that would result in the sum described above equal to unity, since this would imply that the net systematic bias on the dark energy parameter is equal to the statistical constraint on that parameter (recall that the curves in Figures 19 and 20 plot the systematics in units of the statistical uncertainty of the survey). Ideally, of course, the goal of the $\mathrm{P}_\delta(k)$ calibration program is to achieve sufficient precision such that the systematics are well below the level of statistical constraints, and so setting this sum to unity simply provides a guideline for the calibration.

In Fig. 21 we have performed the calculation of the power spectrum precision requirement as a function of $\ell_{\rm max}$. All curves pertain to weak lensing-only experiments[3], with the level of photo-z precision coded according to the legend. The top three curves correspond to the optimistic precision estimate (errors in each bin are added in quadrature), the bottom three

[3]We have not illustrated the significance of including galaxy clustering statistics because it is a relatively minor effect, as evidenced by Fig. 19.

curves to the pessimistic estimate (the absolute value of the error in each bin are added). The requirements plotted in Fig. 21 are set according to the magnitude of the systematics in w_0; the w_a-based requirements are very similar.

The optimistic calculation assumes that the $P_\delta(k)$ errors at different wavenumber are completely independent; the pessimistic calculation assumes that the errors are perfectly correlated. Because the level of correlation between matter power spectrum errors at different wavenumbers will not be known, we stress that for a given $\ell_{\rm max}$ the only way to guarantee that $P_\delta(k)$ errors do not contribute significantly to the dark energy error budget is to attain the level of precision illustrated by the bottom curves.

In Fig. 21, the crossing of the $P_\delta(k)$ precision requirement curves pertaining to different levels of $N_{\rm spec}$ may seem somewhat counterintuitive; one might expect that improving photo-z uncertainty can only lead to more stringent demands on the accuracy of the prediction for the matter power spectrum. This intuitive expectation is supported by a theorem proved Appendix A of Ref. [13], in which the authors demonstrate that the net $\Delta\chi^2$ induced by a systematic error is always reduced by the addition of (unbiased) prior information. However, as shown in Appendix B of the same paper when marginalizing over multiple parameters, $\Delta\chi^2$ *per degree of freedom* may increase. In our case, for large maximum multipoles ($\ell_{\rm max} \gtrsim 1000$) adding prior photo-z information and marginalizing over our 62 photo-z parameters leads to a mild increase in the w_0 bias, a fact which we have traced to a mild difference in the degeneracy between w_0 and Ω_Λ for different values of $N_{\rm spec}$.

From Fig. 20 it is evident that for smaller choices of $\ell_{\rm max}$ dark energy biases are less sensitive to matter power spectrum errors on small scales. For each $\ell_{\rm max}$ there is a maximum wavenumber, $k_{\rm req}^{\rm max}$, such that systematic errors in $P_\delta(k)$ for $k > k_{\rm req}^{\rm max}$ do not produce significant biases in dark energy parameters. We estimate $k_{\rm req}^{\rm max}$ by finding the bin in wavenumber at which the induced w_0 bias becomes less than 10% of the maximum magnitude that δw_0 attains in any bin. Because of the steepness of the scaling of $\delta(w_0)/\sigma(w_0)$ with k on scales smaller than the wavenumber at which the systematics attain their maximum magnitude, we find that our $k_{\rm req}^{\rm max}$ estimations are insensitive to the choice for this percentage. We present our results for $k_{\rm req}^{\rm max}$ as a function of $\ell_{\rm max}$ in Fig. 22.

The results in this section provide a set of concrete benchmarks for the campaign of

numerical simulations designed to calibrate the prediction for the matter power spectrum, as well as a guideline for choosing the maximum multipole that should be included in any cosmic shear analysis. For a given $\ell_{\rm max}$, one uses the results presented in Fig. 21 to estimate the precision with which $\mathrm{P}_\delta(k)$ must be predicted on all scales $k < k_{\rm req}^{\rm max}$, where the $k_{\rm req}^{\rm max}$ estimate appears in Fig. 22. Of course choosing smaller values of $\ell_{\rm max}$ naturally decreases the constraining power of the survey, and so the results we present here can be used to inform the optimal choice for $\ell_{\rm max}$ that balances the need for statistical precision against the threat of matter power spectrum systematics.

4.3.3.3 Halo Model In Fig. 23 we illustrate our results for the propagation of systematic errors in halo concentration parameters through to w_0 (top panel) and $w_{\rm a}$ (bottom panel). Each curve corresponds to a calculation in which a single concentration parameter, either c_0 (red), α (green), or β (blue), is systematically offset upwards of its fiducial value by 10%, while assuming that all of the concentration parameters are known with perfect accuracy and precision. The value of the induced systematic error on $w_0(w_{\rm a})$ has been normalized by the statistical constraints on the parameter at the level of photo-z calibration specified by the horizontal axis value for $N_{\rm spec}$. We propagate systematic errors via Eq. 4.11, as in § 4.3.3.1.

The magnitude of the systematic error induced on dark energy parameters monotonically decreases as $N_{\rm spec}$ increases. The physical interpretation of this trend applies to nearly all of the results presented in this chapter, and so we discuss it in detail in §4.4. Briefly, as photo-z priors are relaxed the cosmological interpretation of the weak lensing signal must rely more heavily on precise knowledge of the matter distribution. In the context of the halo model this implies that errors in halo concentrations have more drastic consequences for dark energy parameter inference at lower values of $N_{\rm spec}$.

Dark energy systematics are also less severe when galaxy clustering information is included; for each halo parameter, and at every level of photo-z calibration, the solid curves are smaller in magnitude than the dashed. This trend is to be expected as its analogues have manifested in previous sections: the additional information available in galaxy clustering statistics mitigates the consequences for cosmology of errors in the matter power spectrum

prediction (see Ref. [198] for the analogous benefit of mitigating photo-z systematics by including galaxy correlations). Notice that there is a proportionally greater mitigation of the systematics at lower values of $N_{\rm spec}$; the interpretation of this observation is somewhat subtle. When photo-z priors are weak, galaxy clustering information plays a more important role in the self-calibration of the photo-z parameters[4], σ_z^i and $z_{\rm bias}^i$. As discussed in the preceding paragraph, systematic errors induced by incorrect predictions of halo concentrations are more severe when information about the photo-z distribution is limited. Therefore, because including galaxy clustering information at low $N_{\rm spec}$ has a greater impact on the statistical constraints than when $N_{\rm spec}$ is very large, there is a concomitantly greater mitigation of the induced systematics associated with halo concentration errors at lower values of $N_{\rm spec}$.

In the limit of very large $N_{\rm spec}$, when uncertainty in the photo-z distribution can be neglected, a 10% error in the mean halo concentration c_0 induces a systematic error on dark energy parameters that is comparable to or worse than the statistical constraints. This estimation is consistent with previous results (Refs. [62; 197]). For halo concentration errors made at other levels, we can estimate the induced systematic error on the inferred value of dark energy parameters via simple linear scaling[5]. For example, as can be seen in Fig. 23, in the limit of very large $N_{\rm spec}$ a 10% misestimation of the parameter c_0 induces a systematic error $\delta w_0 = 1.2\sigma(w_0)$ for a weak lensing-only analysis; thus a 25% error on c_0 induces a $\delta w_0 = [1.2 \times (0.25/0.1)]\,\sigma(w_0) = 3\sigma(w_0)$ systematic. For the sake of concreteness we conclude this section with the following rough guideline that is based on Fig. 23 evaluated at $N_{\rm spec} \approx 10^4$ for a joint analysis: in order to guarantee that dark energy systematics induced by halo concentration errors are kept at or below the statistical constraints, the parameter c_0 must be calibrated to an accuracy of 5% or better, the parameter α to an accuracy of 12% or better, and β to better than 25%.

[4]This is an important observation in itself and has been noted elsewhere in the literature (for example, Ref. [198]). Most of the so-called "complementarity" of weak lensing and galaxy clustering stems from these signals calibrating each other's nuisance parameters.

[5]The stability of the derivatives of our observables with respect to halo concentration parameters over a broad range of numerical step sizes ensures the accuracy of this simple linear scaling.

4.4 DISCUSSION

We have studied the significance of matter power spectrum uncertainty for weak lensing measurements of dark energy. Our results can serve as an updated guideline for the calibration requirements on theoretical predictions of $P_\delta(k)$ and photometric redshift distributions. The photo-z requirements revise those in Ref. [119], who modeled nonlinear evolution using the Peacock & Dodds fitting formula, and were thus overly pessimistic about photo-z calibration. We also revise the requirements for precision in the prediction of $P_\delta(k)$ outlined in Ref. [80], who assumed perfect knowledge of the distribution of photometric redshifts and were thus overly optimistic.

Both of our models for uncertainty in the theoretical prediction for the matter power spectrum have been studied previously. Our second model, in which we allow the value of $P_\delta(k)$ to vary freely about its fiducial values in ten bins of bandpower (see §4.2.1 for a detailed description) is based on the treatment in Ref. [80]. Our results are in good agreement with theirs, where applicable. We have generalized their results by 1) studying the self-calibration limit of $P_\delta(k)$ uncertainty, 2) including galaxy clustering statistics in the set of observables, and 3) by taking into account uncertainty in the distribution of photometric redshifts. Our motivation to treat uncertainty in $P_\delta(k)$ and the photo-z distribution simultaneously comes from results presented in the Appendix of Ref. [64], where the authors showed that the photo-z calibration requirements vary significantly depending on the assumed fiducial model of the $P_\delta(k)$ in the nonlinear regime. This result suggests a nontrivial interplay between the photo-z and matter power spectrum calibration demands.

The cause of this interplay has a simple physical interpretation. In weak lensing, there is a degeneracy between the redshift of a galaxy whose image is distorted and the typical size of the overdensity responsible for most of its lensing: at fixed angular scale, correlations in the image distortions of sources at high redshift are produced (on average) by overdensities that are larger in comoving size than those producing correlations at low-redshift. The more precisely the photo-z distribution is known, the narrower the range of possible wavenumbers contributing to the lensing signal. As priors on photo-z parameters are relaxed, the redshifts of the sources are known with decreasing precision, and so more information about the

power spectrum is required in order to compensate. Thus at lower values of $N_{\rm spec}$, dark energy parameter inference at a fixed level of statistical uncertainty requires more precise knowledge of the matter power spectrum.

Consequences of this basic physical picture appear throughout this chapter. For example, in Figures 19 and 23, appearing in §4.3.3.1 and §4.3.3.3, respectively, the degeneracy between source redshift and length scale manifests as dark energy biases being more severe for the lower values of $N_{\rm spec}$. Another example appears in Figure 18 of §4.3.2.2, in which we can see that the statistical constraints on w_0 and $w_{\rm a}$ degrade more rapidly with uncertainty in halo concentrations when $N_{\rm spec}$ is small relative to larger values of $N_{\rm spec}$. Similarly, the statistical constraints on dark energy discussed in §4.3.2.1 degrade more rapidly as priors are relaxed on $\delta\ln(P_{\rm i})$ at lower values of $N_{\rm spec}$. The nontrivial relationship between uncertainty in $\mathrm{P}_\delta(k)$ and in the distribution of photometric redshifts clearly illustrates that a detailed and accurate study of the calibration requirements on future imaging surveys requires the simultaneous account of these contributions to the dark energy error budget that we present here.

The significance of uncertainty in halo concentrations for the dark energy program has also been studied previously [62; 197]. Again, we have generalized their calculations and our conclusions are in good agreement with their results, where commensurable. In particular, Ref. [197] studied the prospects for future weak lensing surveys to self-calibrate halo concentration parameters while simultaneously constraining dark energy parameters. Even with only very modest prior information on photo-z parameters, we agree with the conclusion in Ref. [197] that the prospect for future imaging surveys to self-calibrate uncertainty in halo concentrations is very promising, especially when galaxy correlation statistics are employed in a joint analysis: the statistical degradation on w_0 from self-calibrating the mean, redshift-zero halo concentration c_0 is less than 6% for $N_{\rm spec} = 5000$ (corresponding roughly to $\Delta\sigma_z/\sigma_z \approx 10^{-2}$). The degradation in the constraints is even milder when the photo-z distribution is more precisely characterized. However, this result is provisional in that it relies on halo concentration being the most significant mode in which the power spectrum is uncertain.

Our halo model-based treatment of $\mathrm{P}_\delta(k)$ uncertainty is well-motivated by Ref. [155], who carried out a suite of numerical cosmological simulations including hydrodynamics with

a variety of energy feedback mechanisms. One of the salient conclusions of Ref. [155] is that the effects of baryonic physics on the matter power spectrum can be well-modeled as an enhancement to the mean concentration of dark matter halos. Thus we chose our first model of uncertainty in $P_\delta(k)$ with the intention to study the requirements of the dark energy program for precision in our ability to predict the effects of baryons on the large-scale distribution of matter. However, recent results from the OverWhelmingly Large Simulation (OWLS) project [156] suggest that an energy feedback mechanism modeling the effects of AGN is necessary to reproduce the characteristics of groups of galaxies [123]. In a recent study based on these results [166], the authors found that a multicomponent halo model with a gas profile that is independent from the dark matter profile can accurately model the power spectra in the OWLS project. We note, however, that not even N-body simulations have achieved the desired precision ($< 0.5\%$) over the full range $0.1\ h\mathrm{Mpc}^{-1} \lesssim k \lesssim 5\ h\mathrm{Mpc}^{-1}$ required of the $P_\delta(k)$ calibration [69]. Nonetheless, these results are intriguing and suggest that a more complicated model than the one we consider here may be necessary to fully encapsulate the baryonic modifications to the matter power spectrum. We leave the development and exploration of such a model as a task for future work.

When the assumption that the halo model accurately characterizes all the gross features of $P_\delta(k)$ is relaxed, self-calibrating the matter power spectrum is very likely to be infeasible. To see this, we turn back to our second, more conservative model of matter power spectrum uncertainty. As shown in Fig. 16, even in the limit of perfectly precise prior knowledge on the photo-z distribution, future surveys are unable to self-calibrate the value of $P_\delta(k)$ to better than 7% on any scale, rendering $P_\delta(k)$ as a dominant component in the error budget and increasing errors on w_0 and w_a by a factor of 3 or more. As discussed in §4.3.3.1, ensuring that dark energy systematics induced by misestimations of $P_\delta(k)$ are kept at or below the level of statistical uncertainty of an LSST- or Euclid-like survey, the theoretical prediction for the matter power spectrum will need to be accurate to at least 0.5% or better on all scales $k \lesssim 5\ h\mathrm{Mpc}^{-1}$. These results reinforce the necessity of an aggressive campaign of numerical cosmological simulations if surveys such as LSST or Euclid are to achieve their potential as dark energy experiments.

To contextualize these findings with the current state-of-the-art in numerical simula-

tions, we first compare this requirement to the results from the Coyote Universe project [68], a suite of nearly 1,000 N-body (gravity only) simulations spanning 38 fiducial wCDM cosmologies. To date, this is the most ambitious campaign of N-body simulations yet performed with the aim to robustly calibrate $\mathrm{P}_\delta(k)$ over the full range of scales relevant to weak lensing. The power spectrum emulator based on their results has recently been completed [106], and in Ref. [69] the authors demonstrate that results from their simulations can be used to model $\mathrm{P}_\delta(k)$ with sub-percent accuracy on all scales $k \lesssim 1\ h\mathrm{Mpc}^{-1}$. In addition to the need to expand the range of scales over which this level of precision has been attained, the Coyote Universe does not account for the nonlinear effects of neutrino mass [1], which have been recently established [16] to introduce percent-level changes to the standard Smith et al. method of prediction. Moreover, Coyote's 1% precision only applies to wCDM cosmologies; extending this level of precision to dynamical dark energy models is an active area of current research on in this field, for example, Refs. [26; 91; 124; 147]. These complications aside, the Coyote Universe project is, by itself, insufficient to completely calibrate the matter power spectrum because N-body simulations neglect the effect that baryonic gas has on $\mathrm{P}_\delta(k)$. Suites of hydrodynamical simulations such as the OWLS project [186] discussed above will be essential contributions to the calibration program. Continued improvement both in N-body and hydrodynamical simulations will clearly be necessary in order to meet the calibration requirements we present here.

4.4.1 Caveats

In updating the photo-z precision requirements for future imaging surveys we have quantified the uncertainty in the photo-z distribution in terms of N_{spec}. We reiterate here an important difference between the meaning of N_{spec} in our forecasts and elsewhere in the literature. In this work, the quantity N_{spec} defines a one-parameter family of priors on the photo-z distribution via Eqs. 4.4 & 4.5. In practice the actual number of galaxies in the calibration sample will likely need to be larger than N_{spec}, for example because it will be challenging to obtain a calibration sample that fairly represents the color space distribution of the galaxies in the imaging survey. Moreover, even if such a representative sample is obtained in a

particular patch of sky, sample variance due to the relatively narrow sky coverage of current and near-future calibration samples has a significant impact on the accuracy of the calibration [40]. We sought to provide general guidelines for a broad range of future imaging surveys, and so we have not attempted to model how these important, survey-specific issues affect the calibration requirements. Instead, in our formulation the photo-z precision requirements are formally specified in terms of the necessary amount of prior knowledge on the photo-z distribution, which in turn is encoded by the parameter $N_{\rm spec}$.

Our modeling of galaxy clustering has several simplifying assumptions that are relevant to the calibration of future imaging surveys. In modeling galaxy bias as a function of redshift only, we have implicitly assumed perfect knowledge of how galaxy bias depends on wavenumber. Uncertainty in the scale-dependence of the galaxy bias is degenerate with uncertainty in $\mathrm{P}_\delta(k)$, and so the improvement in the constraining power of a survey provided by including galaxy clustering will degrade when accounting for uncertainty in the scale-dependence of galaxy bias. However, we have restricted the range of angular scales on which galaxy correlations are exploited so that we probe only very large scales ($\sim 50h^{-1}$Mpc), and so we expect corrections accounting for this scale-dependence are small.

The benefit of galaxy correlations to the photo-z calibration is diluted by lensing magnification bias [13; 160], which induces a spurious correlation between sources that are well separated in redshift space. This effect thus threatens the ability of galaxy cross-correlations to detect and calibrate outliers in the photo-z distribution and will need to be accounted for in order to fully realize the potential of galaxy clustering. We intend to generalize our results to include these effects in future work.

4.5 SUMMARY AND CONCLUSIONS

We have studied the matter power spectrum calibration requirements for future very-wide-area weak lensing surveys such as LSST or Euclid. While our findings apply to all planned imaging surveys designed to use weak lensing to study dark energy, we have phrased our conclusions in terms of these particular surveys because their calibration demands are the

most stringent. Our results generalize previous findings by simultaneously accounting for photometric redshift uncertainty, as well as by studying the significance of galaxy clustering information. We explored two different models for uncertainty in the nonlinear physics of gravitational collapse, which we describe in detail in §4.2.1. In our first model, we assume that the Halo Model accurately predicts the gross shape of $P_\delta(k)$, but that the internal structure of halos is uncertain. In our second, more agnostic model, we allow the value of $P_\delta(k)$ to vary freely about its fiducial value in ten logarithmically-spaced bins of bandpower spanning the range $0.01\ h\mathrm{Mpc}^{-1} \leq k \leq 10\ h\mathrm{Mpc}^{-1}$. We conclude this chapter by providing a brief summary of our primary results.

1. Future imaging surveys will be unable to self-calibrate the value of $P_\delta(k)$ to better than 7% on any scale. This renders infeasible the possibility of completely self-calibrating the theoretical prediction for the matter power spectrum because systematic errors at such levels would induce unacceptably large biases in the inferred value of the dark energy equation of state. Moreover, the marginalized constraints on $P_\delta(k)$ are the tightest at scales $k \approx 0.2\ h\mathrm{Mpc}^{-1}$, nearly an order of magnitude larger in size than where the unmarginalized constraints computed in Ref. [80] attain their minimum, emphasizing the necessity of a precise calibration of $P_\delta(k)$ over the full range of wavenumbers $0.1\ h\mathrm{Mpc}^{-1} \lesssim k \lesssim 5\ h\mathrm{Mpc}^{-1}$.

2. To ensure that systematics are kept at levels comparable to or below the statistical constraints on w_0 and w_a, $P_\delta(k)$ must be accurately predicted to a precision of 0.5% or better on all scales $k \lesssim 5\ h\mathrm{Mpc}^{-1}$ in advance of future weak lensing observations that will be made by LSST or Euclid.

3. The required precision for the $P_\delta(k)$ prediction as well as the scale to which this precision must be attained depend sensitively on ℓ_max, the maximum multipole used in the cosmic shear analysis. Figures 21 and 22 together provide a concrete guideline that can be used to directly inform the optimal choice for ℓ_max that balances the need for statistical precision against the threat of matter power spectrum systematics.

4. In keeping with the results in the Appendix of Ref. [64], we find that the photo-z calibration requirements are less stringent by a factor of ~ 3 when the nonlinear evolution

of $P_\delta(k)$ is modeled with the Smith et al. fitting formula relative to Peacock & Dodds, significantly relaxing the demands for photo-z precision that appear in Ref. [119].

5. Dark energy constraints degrade $\sim 40\%$ more slowly with photo-z uncertainty when including galaxy correlations in a joint analysis with weak lensing, even when the clustering information is restricted to degree-scales and with coarse tomographic redshift binning so that Baryonic Acoustic Oscillation features are not resolved.

6. Including galaxy clustering statistics (again, even when Baryon Acoustic Oscillation information is neglected) also significantly relaxes the calibration requirements and mitigates the severity of systematic errors induced by erroneous predictions for $P_\delta(k)$, especially when prior information on the photo-z distribution is weak. Dark energy systematics can be reduced by up to 50% by including galaxy clustering information; the statistical constraints on w_0 and w_a can degrade $2-5$ times more rapidly with $P_\delta(k)$ uncertainty when galaxy correlations are neglected.

7. The redshift-zero, mean halo concentration, c_0, must be accurately predicted with a precision of 5% or better to keep systematics in dark energy parameters below the level of statistical constraints. If internal halo structure is the dominant mode of $P_\delta(k)$ uncertainty, then the prospect for self-calibrating c_0 are quite promising, as this would only degrade the dark energy constraints by $5-10\%$.

8. The matter power spectrum calibration requirements are more stringent when the distribution of photometric redshifts is known with less precision. This effect is due to a degeneracy between source redshift and lens size, and is the chief motivation for a simultaneous account of these sources of uncertainty. We find that the constraints on w_0 and w_a degrade $2-3$ times more rapidly with $P_\delta(k)$ uncertainty for spectroscopic calibration samples with the statistical equivalent of $N_{\rm spec} \approx 5000$ relative to $N_{\rm spec} \approx 10^5$.

9. The requirements for the precision with which $P_\delta(k)$ need be predicted are, in general, less stringent for DES than for LSST by a factor of a few. To ensure that matter power spectrum systematics do not contribute significantly to the dark energy error budget for DES, we find that if correlations up to a maximum multipole of $\ell_{\rm max} = 3000$ are used in the lensing analysis then $P_\delta(k)$ will need to be calibrated to an accuracy of 2% or better on scales $k \lesssim 5\ h{\rm Mpc}^{-1}$. The DES requirements scale with $\ell_{\rm max}$ in a similar fashion to

the scaling of the LSST requirements summarized in Fig. 22. The difference between the requirements is driven by the relative depth of these two surveys: the constraining power on dark energy provided by cosmic shear measurements derives chiefly from small-scale correlations ($k \gtrsim 1h\mathrm{Mpc}^{-1}$) where shot noise is most significant. The shallower depth of DES ($z_{\mathrm{med}} = 0.7$, $N_{\mathrm{A}} = 15$ gal/arcmin^2) results in these modes being less informative, and so DES suffers less from uncertainty in small-scale information.

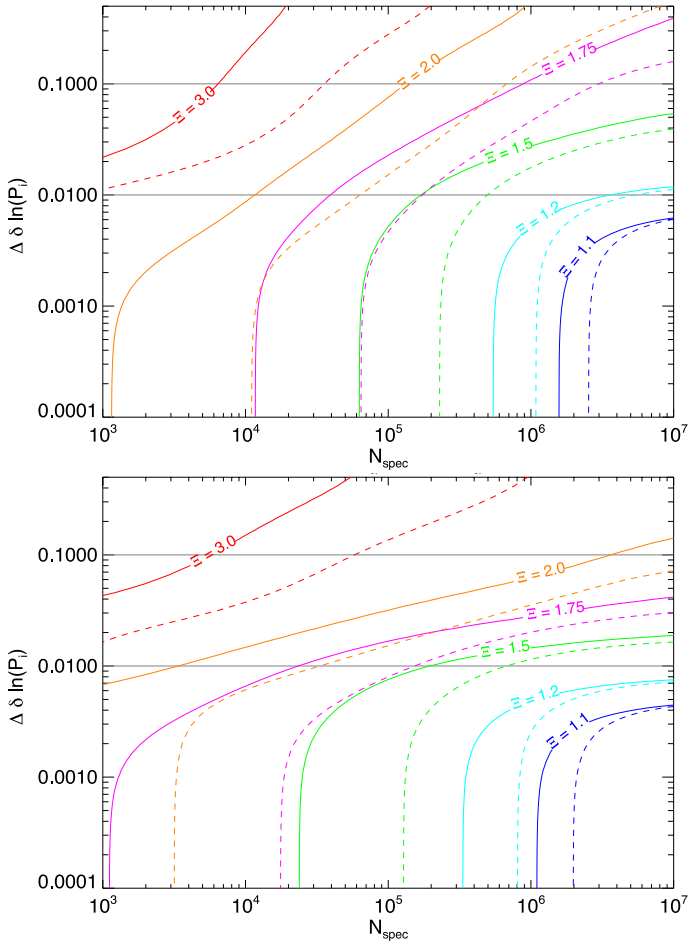

Figure 17: Contours of degradation in w_0 and w_a constraints appear in the top and bottom panels, respectively. The degradation is quantified by $\Xi \equiv \sigma(w_0, w_a)/\sigma^{\rm perf}(w_0, w_a)$, where $\sigma^{\rm perf}$ is the level of statistical uncertainty of an LSST- or Euclid-like survey in the limit of perfect certainty on the photo-z distribution and the nonlinear evolution of $P_\delta(k)$. The precision in the calibration of the matter power spectrum $P_\delta(k)$ appears on the vertical axes, while the priors on the photo-z distribution, parameterized by $N_{\rm spec}$, appear on the horizontal axes. Prior information about the functions $z^{\rm bias}$ and σ_z, which govern the uncertainty in the photo-z distribution $P(z^{ph}|z^{sp})$, is distributed uniformly in redshift according to $\Delta z^{\rm bias} = \sqrt{2}\Delta\sigma_z = \sigma_z/\sqrt{N_{\rm spec}}$. Priors on the $\delta\ln(P_i)$ parameters are distributed uniformly in among bandpowers in $\log(k)$. Dashed curves pertain to a survey using only weak lensing information, solid curves to a joint analysis that includes galaxy clustering. The gray, horizontal lines roughly bound the range of matter power spectrum uncertainty that is attainable in advance of these surveys.

Table 2: Baseline Constraints

Observables	$\sigma(w_0)$	$\sigma(w_{\rm a})$	$\sigma(w_{\rm piv})$
Weak Lensing Only	0.071	0.22	0.022
Joint Analysis	0.058	0.19	0.018

NOTES.— Column (1) specifies whether or not observables employing galaxy clustering ($P_{\rm gg}$ and $P_{\kappa{\rm g}}$) were used in the calculation. Columns (2), (3), and (4) give the statistical constraints on w_0, $w_{\rm a}$, and $w_{\rm p}$, respectively, that can be obtained by future very-wide-area surveys such as LSST or Euclid. Note that these constraints account for statistical errors only and so represent the optimistic limit of achievable dark energy constraints for a survey with these characteristics.

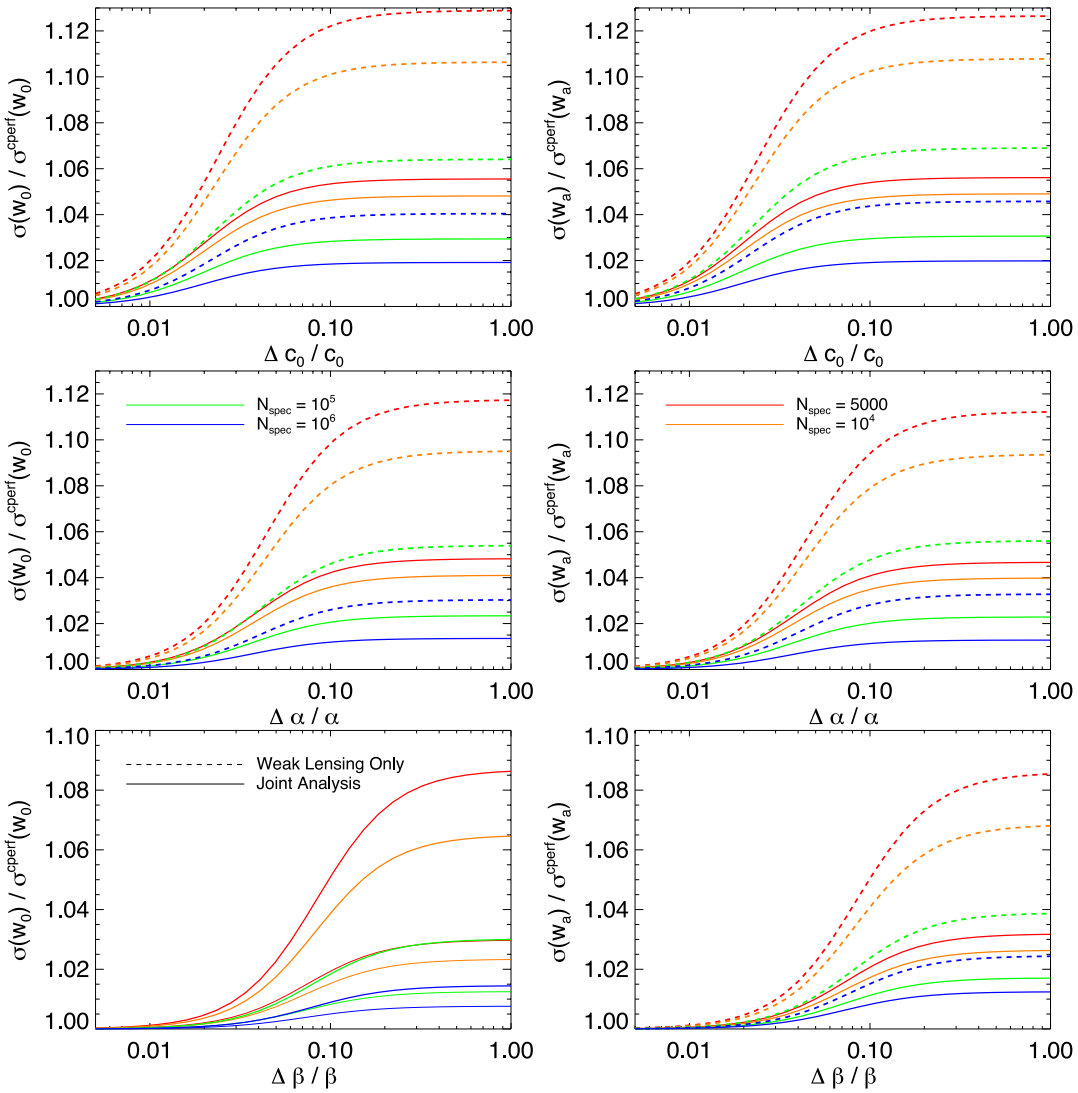

Figure 18: Plot of the degradation in dark energy constraints as a function of priors on halo concentration parameters, defined by $c(m,z) = c_0[m/m_{*,0}]^{\alpha}(1+z)^{\beta}$. Degradation of the constraints on w_0 are shown in the *left* panels, while degradation of w_a constraints are shown in the *right* panels. The *top* panels show degradation as a function of the fractional prior uncertainty in the normalization of the mass concentration relation c_0. The *middle* row of panels shows degradation as a function of the power law index describing the dependence of concentration on mass, α, and the *bottom* panels show the degradation as a function of the power-law index describing the dependence of concentration on redshift, β. Different levels of priors on photo-z parameters are color-coded as labeled in the middle panels.

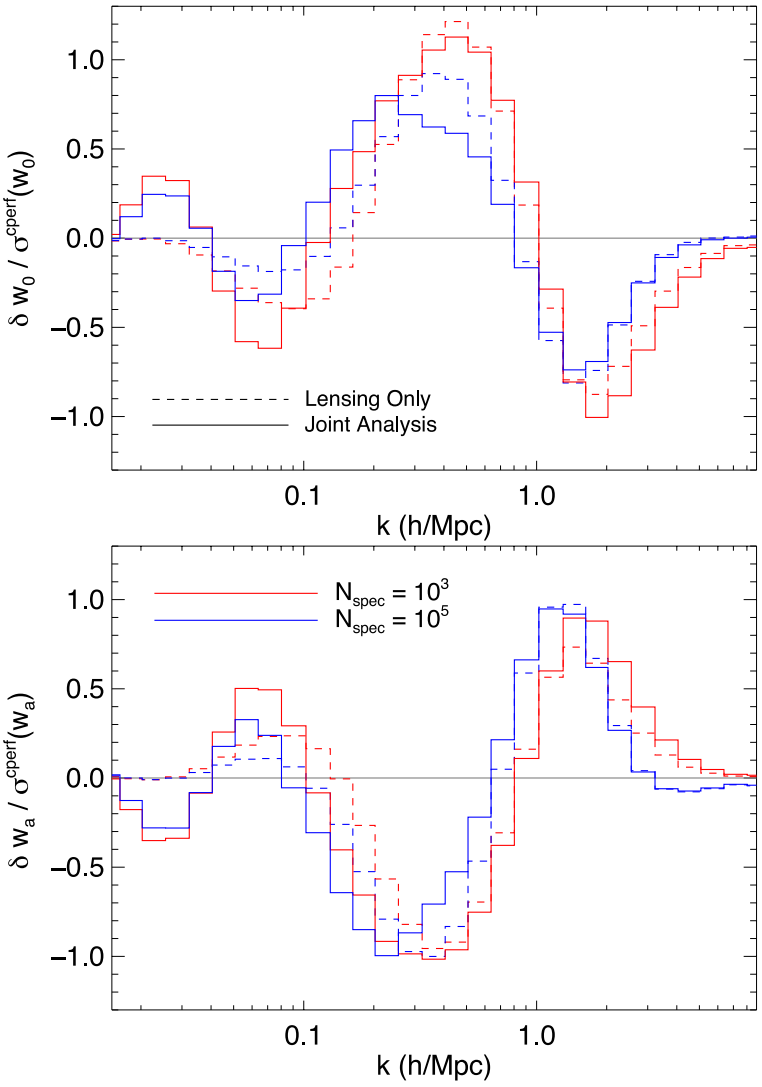

Figure 19: Plot of the systematic error in the inferred value of w_0 (top panel) and $w_{\rm a}$ (bottom panel) induced by a systematic misestimation of the matter power spectrum $\mathrm{P}_\delta(k)$ over a small range of wavenumbers. The value of the induced systematic is scaled in units of the statistical uncertainty of an LSST- or Euclid-like survey, and is plotted against the scale at which $\mathrm{P}_\delta(k)$ has been incorrectly predicted by 5% over a range of wavenumbers of width $\Delta k/k = 0.1$. See the text for a detailed description of how these results can be rescaled for $\mathrm{P}_\delta(k)$ errors spanning a range of wavenumbers of a different width. Results pertaining to an analysis that only uses weak lensing information appear as dashed curves, a joint analysis that includes galaxy clustering as solid curves. We show results for two different levels of photo-z calibration, with the red curves corresponding to $N_{\rm spec} = 10^3$, and the blue curves to $N_{\rm spec} = 10^5$.

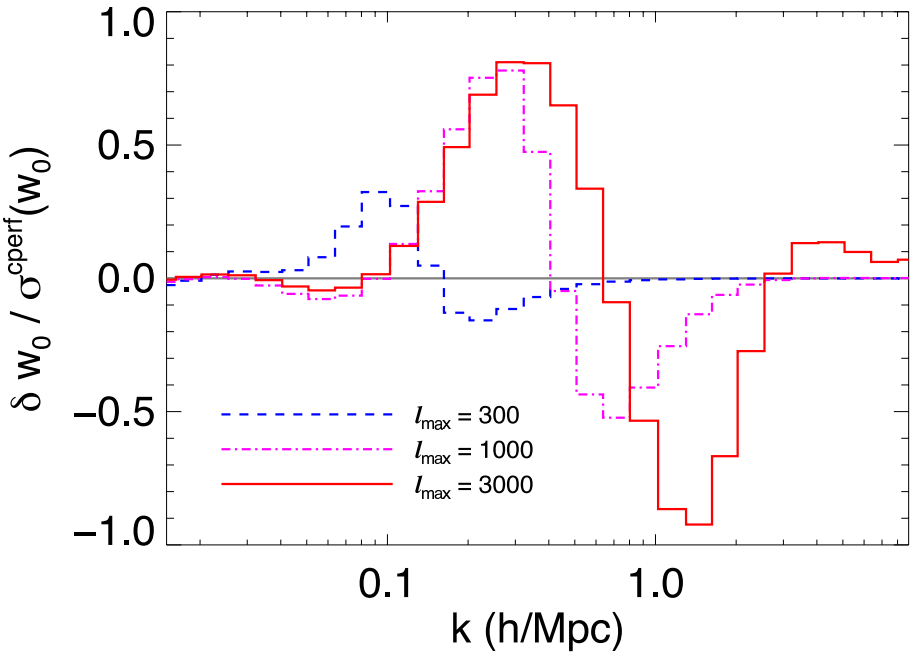

Figure 20: Plot of the systematic error in the inferred value of w_0 (top panel) induced by a systematic misestimation of the matter power spectrum $\mathrm{P}_\delta(k)$. This plot is identical to Figure 5 except the curves illustrate weak lensing-only results in the limit of perfect core calibration ($N_{\rm spec} \to \infty$) but with different choices for the maximum multipole. Naturally, the sensitivity of the experiment to small-scale $\mathrm{P}_\delta(k)$ errors decreases as $\ell_{\rm max}$ decreases.

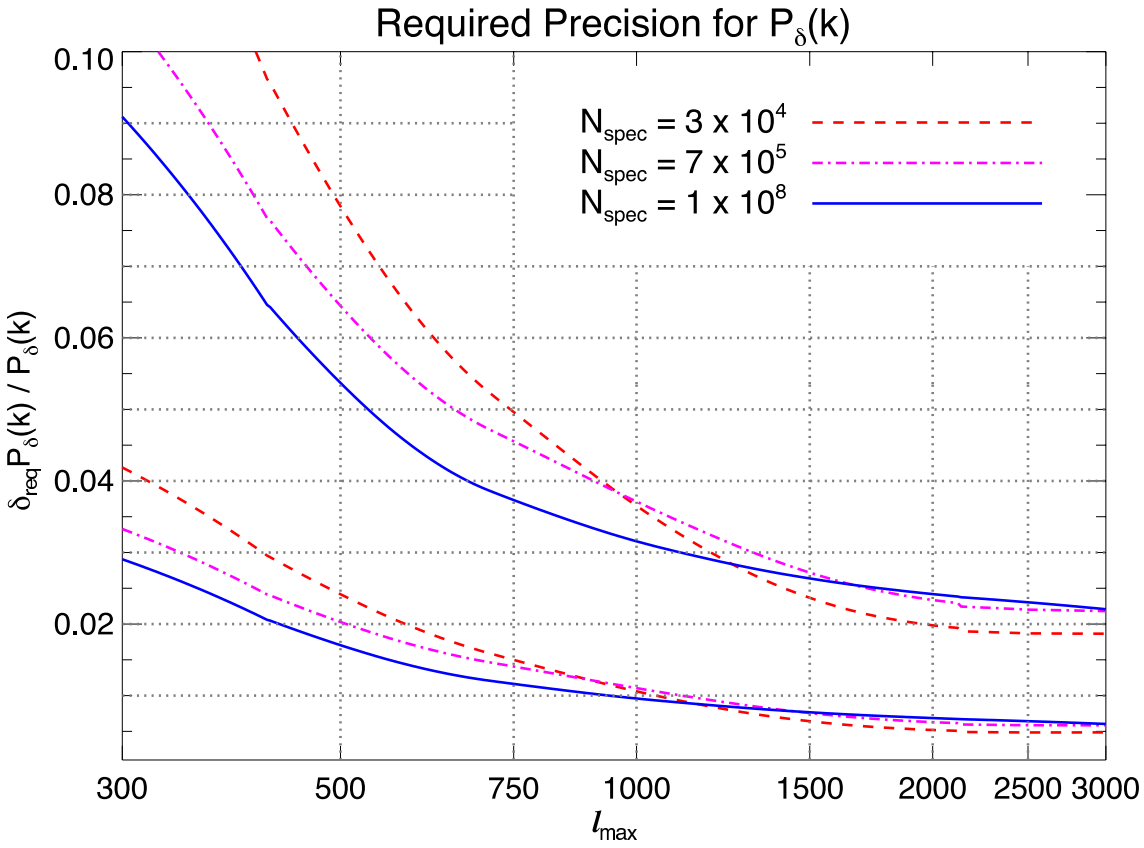

Figure 21: Plot of the required precision with which the matter power spectrum must be predicted as a function of $\ell_{\rm max}$. The top curves represent an optimistic estimate, the bottom curves a pessimistic estimate. All curves represent results for a cosmic shear-only analysis; calculations for different levels of photo-z precision are color-coded according to the legend. To guarantee that matter power spectrum errors do not contribute significantly to the dark energy error budget, the precision in the prediction for $P_\delta(k)$ must reach the level illustrated by the bottom curves. One uses the results in Figure 8 to estimate the wavenumber to which the level of precision plotted here must be attained. See text for details concerning the calculation of these estimates.

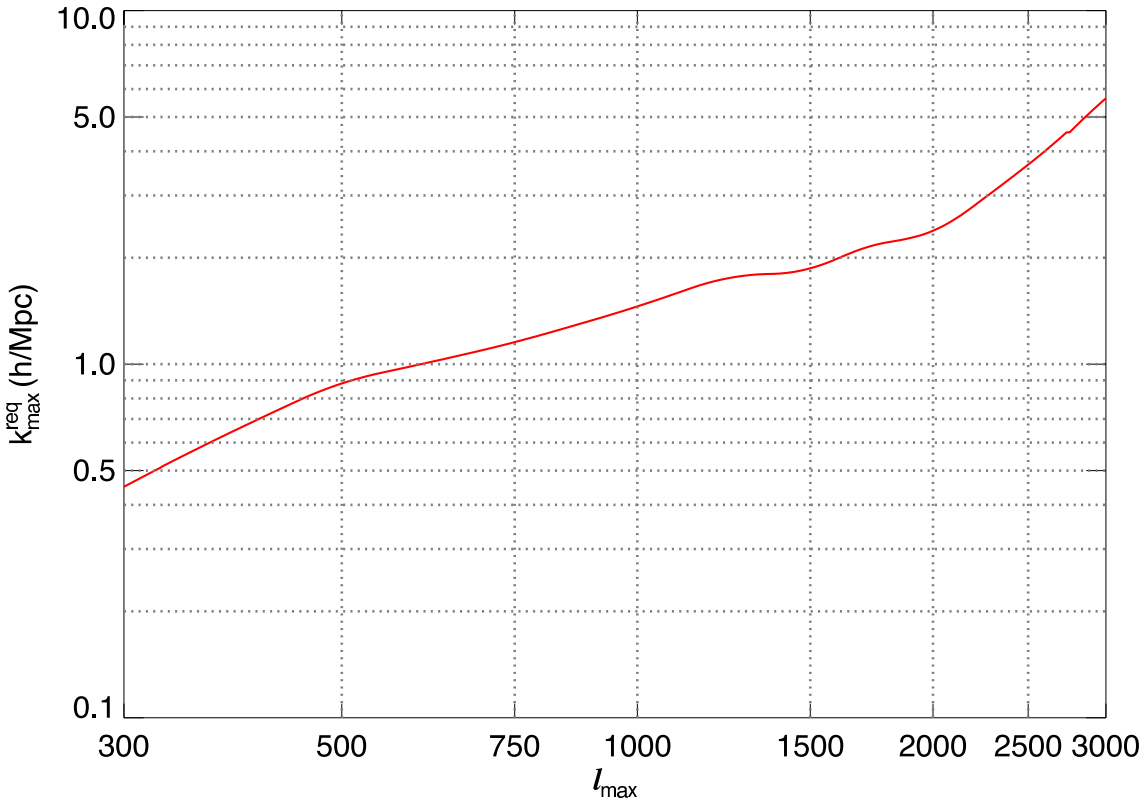

Figure 22: Plot of $k_{\rm req}^{\rm max}$, the wavenumber to which the matter power spectrum must be predicted, as a function of $\ell_{\rm max}$. The quantity $k_{\rm req}^{\rm max}$ is insensitive to the level of photo-z precision, and so the same $k_{\rm req}^{\rm max}$ pertains to all values of $N_{\rm spec}$. One uses the results of Figure 7 to estimate the level of precision in the prediction for $\mathrm{P}_\delta(k)$ that needs to be attained for all $k < k_{\rm req}^{\rm max}$.

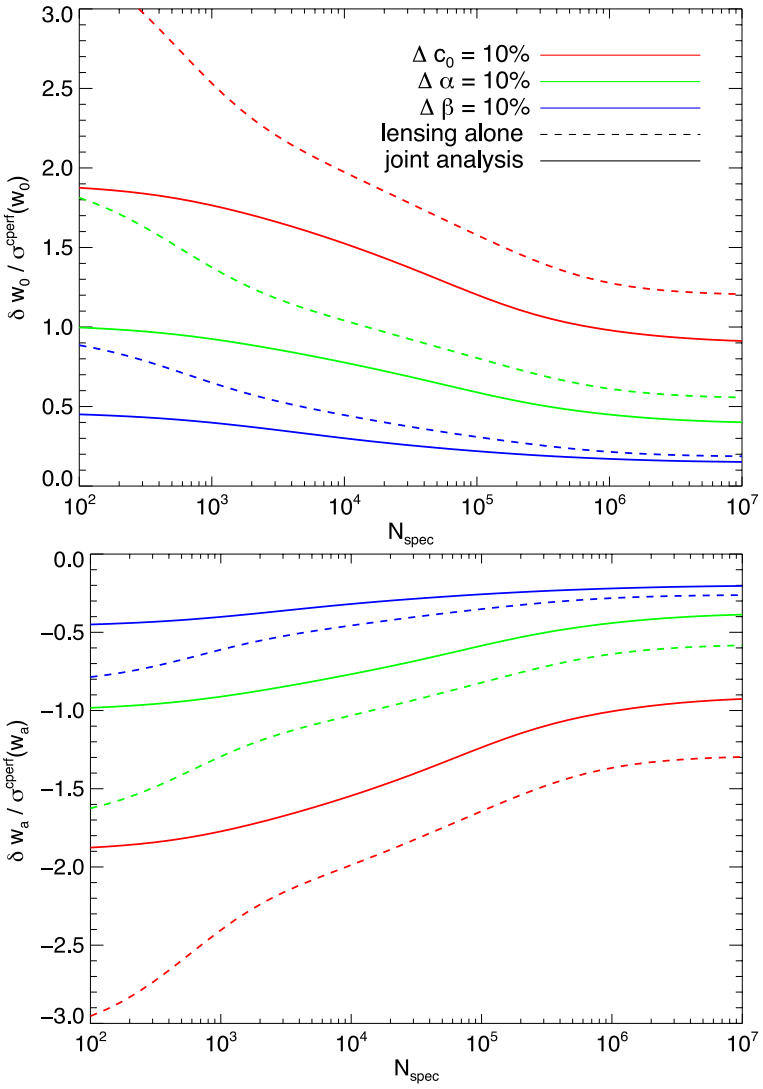

Figure 23: In the *top* panel we plot the systematic error on w_0 against $N_{\rm spec}$, the quantity encoding the precision with which the photo-z distribution has been calibrated. The same quantity is plotted for $w_{\rm a}$ in the *bottom* panel. Dark energy systematics $\delta(w_0, w_{\rm a})$ have been scaled by $\sigma^{\rm cperf}(w_0, w_{\rm a})$, the statistical constraints on the parameter at the level of $N_{\rm spec}$ given by the horizontal axis value of $N_{\rm spec}$. The systematic errors are induced by 10% misestimations of our three halo concentration parameters c_0, α, and β. Results for each parameter are color-coded according to the legend in the *top* panel. The concentration parameters are defined by $c(m,z) = c_0[m/m_{*,0}]^{\alpha}(1+z)^{\beta}$. Dashed curves correspond to results for an analysis using weak lensing information only, solid curves an analysis employing galaxy clustering information together with weak lensing in a joint analysis.

5.0 TESTING THE ORIGIN OF THE CMB LARGE-ANGLE CORRELATION DEFICIT WITH A GALAXY IMAGING SURVEY

5.1 INTRODUCTION

The consistency of the ΛCDM model of cosmology with the cosmic microwave background (CMB) data observed by the Wilkinson Microwave Anisotropy Probe (WMAP) is one of the crowning achievements of twentieth-century cosmology. Indeed, these observations were among the key results that led to the widespread acceptance of the "concordance model" of cosmology.

Despite the remarkable agreement between ΛCDM predictions and the WMAP data, several anomalies on the largest angular scales have persisted (see Refs. [36] and [12] for recent reviews). Arguably the most troubling anomaly is the near-total lack of correlation in the temperature anisotropy distribution on large scales: the CMB temperature autocorrelation function $C(\theta)$ as measured by WMAP is near zero for angular scales above 60 degrees. This puzzling observation first appeared in the Cosmic Background Explorer (COBE) data [172] before detection at much higher significance by WMAP. Quantifying how unusual the large-angle correlation deficit is must be done with care because none of the statistics describing the anomaly itself were among the estimators proposed by the WMAP team prior to undertaking their analysis; that is, analysis of the correlation deficit is complicated by its *a posteriori* observation. Recent estimates of the degree to which this anomaly is in conflict with ΛCDM vary significantly (e.g., Refs. [34; 35; 48; 146]) and especially depend on the treatment of the Galactic region of the microwave sky. For example, the analysis in Ref. [48] employs a reconstruction technique to first generate a full-sky map and then use an all-sky estimator of $\mathcal{C}(\theta)$, concluding that the lack of CMB temperature correlation at large angles is unlikely at

roughly the 95% level. However, in a recent analysis [37] it was demonstrated that leakage of information from the masked region of the sky can lead to biases in the low multipoles computed from a reconstructed full-sky map (see also Ref. [53]). In an alternative approach to Ref. [48], the analysis in Ref. [35] uses a pixel-based estimator of $\mathcal{C}(\theta)$ that is constructed strictly from a cut-sky map to conclude that the possibility of the large-angle correlation deficit being a statistical fluke is unlikely at the level of 99.975%. The appropriate treatment of the Galactic plane remains an active area of research and a widespread consensus has not yet been reached, though this issue is central to the analysis of the low-multipole anomalies. Nonetheless, it is quite clear that the correlation deficit is not likely to be explained away as a simple systematic error: as demonstrated in Ref. [23], this feature cannot be accounted for with a statistically independent contaminant such as an undiagnosed foreground, as such a contaminant would only contribute additional large-scale power, thereby exacerbating the anomaly.

One possible explanation that would naturally produce the observed deficit in the correlation function at large angles is that the primordial power spectrum generated in the early universe is suppressed on comoving scales comparable to the size of the horizon at the time of recombination. The persistence of the large-angle anomaly, coupled with the possibility of accounting for it with suppressed primordial power, has motivated many studies of possible mechanisms for the suppression in the context of inflation, e.g., Refs. [18; 20; 31; 55; 85; 88; 104; 105; 110; 145; 170; 194]. Thus the deficit in $C(\theta)$ at large angles may be an indication that the simplest and most widely accepted models of inflation require revision.

The distinct advantage of using the CMB to probe very large scales is that the photons which free stream to Earth from the surface of last scattering have the potential to transmit information to us from the highest redshift in the visible universe. Thus, in addition to CMB temperature, the polarization signal in the CMB can also potentially be exploited to provide useful information about the distribution of matter on the largest scales [126; 128]. Additionally, the scattering of CMB photons off of galaxy clusters induces a polarization signal that may be possible to exploit to probe very large scales [10; 22; 94; 167]. However, if primordial power in the early universe was in fact suppressed on large scales, then the signature of this suppression should in principle also be imprinted in the distribution of

large-scale structure at low redshift. Confirmation of this signal in, for example, galaxy clustering statistics would independently test the authenticity of the low-power anomaly as a genuine feature of our cosmology.

In Ref. [57], the authors demonstrate that a forthcoming redshift survey such as BigBOSS [158] has the potential to constrain viable models of primordial power suppression (see also Ref. [96]). In this work, we attempt to answer a related question: what is the potential for future galaxy imaging surveys such as the Large Synoptic Survey Telescope (LSST) [3] or Euclid [150] to test the authenticity of the large-angle CMB temperature correlation deficit? On the one hand, photometric redshift (hereafter *photo-z*) uncertainty restricts the number of radial modes that will be available to an imaging survey relative to a data set with spectroscopic redshifts. However, the much larger surface density of sources in an imaging survey dramatically reduces errors due to shot noise at high redshift and allows for the possibility of utilizing the cosmic shear signal in addition to galaxy clustering. Thus through a joint analysis of cosmic shear and galaxy clustering, an imaging survey such as LSST or Euclid may be able to provide independent tests of the large-angle CMB anomalies in the near future.

This chapter is organized as follows. In §5.2 we describe our methods for modeling primordial power suppression and assessing its detectability in the distribution of large-scale structure. In §5.3 we present our results, and we discuss our conclusions in §5.4.

5.2 METHODS

5.2.1 Primordial Power Suppression

To model the suppression of power on large scales, we modify the dimensionless curvature power spectrum $\Delta^2_R(k) \equiv k^3 P_R(k)/2\pi^2$ by introducing a prefactor that encodes the exponential suppression, $\Delta^2_R(k) \to \Delta^2_R(k)\mathcal{S}(k)$, with

$$\mathcal{S}(k) \equiv 1 - \beta \exp\left[-(k/k_c)^\alpha\right]. \tag{5.1}$$

In Eq. 5.1, k_c governs the comoving scale at which suppression in Δ^2_R becomes significant, β the maximum fractional amount of suppression, and α the rapidity with which the suppression approaches its maximal effect. We model Δ^2_R as in Eq. 5.1 not to advocate a particular alternative physical theory of the early universe, but rather so that we can have a simple model for the deficit in $C(\theta)$ at large angles whose consequences can then be explored. As for the particular functional form we choose for $\mathcal{S}(k)$, our motivation is twofold. First, the authors in Ref. [57] demonstrated that ΛCDM cosmology with appropriately tuned exponential suppression of primordial power is a better fit to WMAP data than standard (unsuppressed) ΛCDM.[1] In addition to this phenomenological motivation, Eq. 5.1 has been explored previously in the forecasting literature (in particular, Refs. [31; 57; 126]) and so adopting this model facilitates a comparison of our calculations with existing results.

The convention in the literature that has arisen to describe the large-angle anomaly is the so-called $S_{1/2}$ statistic, defined as

$$S_{1/2} \equiv \int_{-1}^{1/2} \mathrm{d}(\cos\theta)\mathcal{C}(\theta). \tag{5.2}$$

As shown in Ref. [57], the value of the log of the cutoff parameter favored by the $S_{1/2}$ statistic alone is $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) = -2.7$, whereas a joint fit of both the C_ℓ's and $S_{1/2}$ favors models with $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) = -3.3$. Of course the likelihood of a given suppression model depends on the statistic used in the quantification, but regardless of this choice the likelihood is most sensitive to the cutoff scale, k_c, so unless otherwise stated we choose parameter values $\alpha = 3$ and $\beta = 1$ as our fiducial model of suppression and treat k_c as a free parameter.

5.2.2 Power Spectra

To assess the detectability of primordial power suppression in an imaging survey, our basic observables will be two-dimensional projected power spectra. The power spectrum $\mathcal{P}^{s_i s_j}(k, z)$ associated with the correlation function of a pair of three-dimensional scalar fields, $\mathrm{s_i}$ and

[1] Of course this should not be surprising: the model for primordial power spectrum defined by Eq. 5.1 has been specifically constructed to improve the likelihood at large angles.

$\mathrm{s_j}$, can be related to its two-dimensional projected power spectrum, $P^{\mathrm{x_i x_j}}(\ell)$, via the Limber approximation:

$$P^{\mathrm{x_i x_j}}(\ell) = \int dz \frac{\mathrm{W}_i(z)\mathrm{W}_j(z)}{\mathrm{D}_\mathrm{A}^2(z)H(z)} \mathcal{P}^{\mathrm{s_i s_j}}(k = \ell/D_\mathrm{A}(z), z). \tag{5.3}$$

In Eq. 5.3, valid for $\ell \gtrsim 10$, $\mathrm{x_i}$ and $\mathrm{x_j}$ are the 2-D fields that we observe as projections of the 3-D fields, $\mathrm{s_i}$ and $\mathrm{s_j}$, respectively. The angular diameter distance function is denoted by D_A, and $H(z)$ is the Hubble expansion parameter. The 2-D projected fields are related to the 3-D source fields through an integral over an appropriate weight function, $\mathrm{W}_i(z)$, associated with the observable of interest:

$$\mathrm{x_i}(\hat{\mathbf{n}}) = \int dz \mathrm{W}_i(z)\mathrm{s_i}(\hat{\mathbf{n}}, z). \tag{5.4}$$

For galaxy fluctuations, the weight function is simply the redshift distribution[2] of galaxies in the i^{th} tomographic bin, $n_i(z)$, times the Hubble expansion parameter:

$$\mathrm{W}_\mathrm{i}^g(z) = H(z)n_i(z). \tag{5.5}$$

For fluctuations in cosmic shear (convergence), the weight function is given by

$$W_\mathrm{i}^\kappa(z) = \frac{3}{2}H_0^2(1+z)\Omega_\mathrm{m}\mathrm{D}_\mathrm{A}(z)\int_z^\infty dz' \frac{D_\mathrm{A}(z,z')}{D_\mathrm{A}(z')}n_\mathrm{i}(z'), \tag{5.6}$$

where $D_\mathrm{A}(z, z')$ is the angular diameter distance between z and z'.

In Eq. 5.3, the three-dimensional power spectrum of the scalar field sourcing cosmic shear is the matter power spectrum $\mathcal{P}^{\mathrm{s_i s_j}}(k, z) = \mathcal{P}^{\delta_\mathrm{i}\delta_\mathrm{j}}(k, z)$, whereas for galaxy-galaxy correlations the source power is the 3-D galaxy power spectrum $\mathcal{P}^{\mathrm{s_i s_j}}(k, z) = \mathcal{P}^{g_\mathrm{i} g_\mathrm{j}}(k, z)$. Since we will be interested in the galaxy distribution on very large scales, it will suffice for our purposes to relate the galaxy overdensity to the matter overdensity through a simple linear bias, $g_\mathrm{i} = b_\mathrm{i}\delta_\mathrm{i}$, so that the galaxy power spectrum is related to the matter spectrum as $\mathcal{P}^{g_\mathrm{i} g_\mathrm{j}}(k, z) = b_\mathrm{i} b_\mathrm{j}\mathcal{P}^{\delta_\mathrm{i}\delta_\mathrm{j}}(k, z)$. In all our calculations we use an independent galaxy bias function in each

[2] The redshift distribution of galaxies, denoted as $n(\theta, \phi, z)$, is not to be confused with the three-dimensional number density of galaxies, $n_\mathrm{V}(\vec{\mathbf{q}}, \chi(z))$, with χ the comoving radial coordinate and $\vec{\mathbf{q}}$ the transverse (2-D) coordinate. The two distributions are simply related to each other as $n(\theta, \phi, z)\mathrm{d}\Omega\mathrm{d}z = n_\mathrm{V}\mathrm{d}q^2\mathrm{d}\chi$, or $n(\theta, \phi, z) = \frac{\mathrm{D}_\mathrm{A}^2(z)}{H(z)}n_\mathrm{V}(\vec{\mathbf{q}}, \chi)$.

tomographic bin and allow each bias function to vary freely about a fiducial value of $b_{\rm i}(z) = 1$. As the bias of most galaxy samples typically increases with redshift this simple prescription is nominally conservative, although in practice we find that our results concerning power suppression are insensitive to the fiducial bias function as well as the number of galaxy bias functions: the influence of scale-independent galaxy bias on our observables does not resemble the effect of a cutoff in the primordial power spectrum.

Above and throughout, lower-case Latin indices label the tomographic redshift bin of the sources. In principle, the redshift distribution of the galaxies used for the galaxy power spectra need not be the same as that used for cosmic shear, but for simplicity we use the same underlying distribution for both so that the chief difference between the galaxy-galaxy power spectrum $P_{g_{\rm i}g_{\rm j}}$, the shear-shear power spectrum $P_{\kappa_{\rm i}\kappa_{\rm j}}$, and the cross-spectrum $P_{\kappa_{\rm i}g_{\rm j}}$, is the form of the weight functions.

As the models of primordial power suppression we study primarily affect very large scales, we will need to relax the Limber approximation in order to accurately predict the power at low ℓ. In this case,

$$P^{{\rm x_i x_j}}(\ell) = \frac{2}{\pi}\int {\rm d}(\ln k) k^3 P^{\delta_{\rm i}\delta_{\rm j}}(k)\mathcal{I}_{\rm x_i}(k)\mathcal{I}_{\rm x_j}(k), \tag{5.7}$$

where

$$\mathcal{I}_{g_{\rm i}}(k) = \int {\rm d}z \frac{{\rm W}_{\rm i}^{g}(z)}{H(z)} b_{\rm i}(z){\rm D_m}(z) j_\ell(k\chi(z)),$$

$$\mathcal{I}_{\kappa_{\rm i}}(k) = \int {\rm d}z \frac{{\rm W}_{\rm i}^{\kappa}(z)}{H(z)} {\rm D_m}(z) j_\ell(k\chi(z)),$$

with ${\rm D_m}(z) \equiv \delta_{\rm m}(z)/\delta_{\rm m}(z=0)$ denoting the growth function, $\chi(z)$ the comoving distance, $b_{\rm i}$ the linear galaxy bias parameter, and $j_\ell(x)$ the spherical Bessel functions. See Ref. [113] for a recent, rigorous derivation of Eq. 5.7.

We model the underlying redshift distribution as $n(z) \propto z^2 \exp -(z/z_0)$ where the normalization is fixed so that $\int_0^\infty {\rm d}z n(z) = N_{\rm A}$, the surface density of sources in the survey. As we will chiefly be interested in predictions for a very-wide-area photometric survey such as LSST or Euclid, unless explicitly stated otherwise we set $N_{\rm A} = 30$ gal/arcmin^2 and $z_0 = 0.34$ so that the median redshift is unity. We follow the treatment in Ref. [119] and relate the tomographically binned galaxy distributions $n_i(z)$ to the underlying redshift distribution

according to

$$n_i(z) = n(z) \int_{z^i_{\rm low}}^{z^i_{\rm high}} {\rm d}z^{\rm ph} P(z^{ph}|z^{sp}),$$

where $z^i_{\rm low}$ and $z^i_{\rm high}$ are the boundaries of the $i^{\rm th}$ tomographic redshift bin. Photo-z uncertainty is controlled by the function $P(z^{ph}|z^{sp})$, a Gaussian at each redshift:

$$P(z^{ph}|z^{sp}) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{(z - z^{\rm ph} - z^{\rm bias})^2}{2\sigma_z^2}\right]. \tag{5.8}$$

The quantities σ_z and $z^{\rm bias}$ are themselves functions of redshift; we model the evolution of the spread as $\sigma_z = 0.05(1+z)$ and set $z^{\rm bias} = 0$ at all redshifts.

For a survey with its galaxies divided into N_g redshift bins used to measure the galaxy correlation function, and N_s bins for the galaxies used to measure cosmic shear, there will be $N_g(N_g+1)/2$ distinct 2-D galaxy-galaxy power spectra $P_{g_{\rm i}g_{\rm j}}$, $N_s(N_s+1)/2$ distinct shear-shear power spectra $P_{\kappa_{\rm i}\kappa_{\rm j}}$, and N_sN_g distinct cross-spectra $P_{\kappa_{\rm i}g_{\rm j}}$.

5.2.3 Covariance

The covariance between a pair of observables, $P^{\rm x_i x_j}$ and $P^{\rm x_m x_n}$, is quantified by the covariance matrix

$$\mathbf{C}[P^{\rm x_i x_j}, P^{\rm x_m x_n}](\ell) = \tilde{P}^{\rm x_i x_m}(\ell)\tilde{P}^{\rm x_j x_n}(\ell) + \tilde{P}^{\rm x_i x_n}(\ell)\tilde{P}^{\rm x_j x_m}(\ell). \tag{5.9}$$

For the case of either galaxy-galaxy or shear-shear, the observed power spectra $\tilde{P}^{\rm x_i x_j}$ have contributions from signal and shot noise,

$$\tilde{P}^{\rm x_i x_j}(\ell) = P^{\rm x_i x_j}(\ell) + {\rm N}^{\rm x_i x_j},$$

where ${\rm N}^{g_{\rm i}g_{\rm j}} = \delta_{\rm ij} N_{\rm i}^{\rm A}$ is the shot noise term for galaxy-galaxy spectra, with $N_{\rm i}^{\rm A}$ denoting the surface density of sources, and ${\rm N}^{\kappa_{\rm i}\kappa_{\rm j}} = \delta_{\rm ij}\gamma_{\rm int}^2 N_{\rm i}^{\rm A}$ is the shear-shear shot noise term. We calculate the observed cross-spectra $\tilde{P}^{\kappa_{\rm i}g_{\rm j}}$ without a contribution from shot noise, so that $\tilde{P}^{\kappa_{\rm i}g_{\rm j}} = P^{\kappa_{\rm i}g_{\rm j}}$. We follow convention and set the intrinsic galaxy shape noise $\gamma_{\rm int}^2 = 0.2$ and absorb differences in shape noise between different observations into the surface density of sources.

5.2.4 Forecasting

We quantify the detectability of a primordial power suppression model with a set of two-dimensional power spectra $P^{\mathrm{x_i x_j}}$ in two distinct ways. First, we compute the χ^2 difference between suppressed and unsuppressed power spectra:

$$\Delta\chi^2 = \sum_{\mathrm{i,j,m,n},\ell} f_{\mathrm{sky}}(2\ell+1)\Delta P_{\mathrm{x_i x_j}}(\ell)\mathbf{C}^{-1}[P_{\mathrm{x_i x_j}}, P_{\mathrm{x_m x_n}}](\ell)\Delta P_{\mathrm{x_m x_n}}(\ell), \tag{5.10}$$

where $\mathbf{C}^{-1}[P_{\mathrm{x_i x_j}}, P_{\mathrm{x_m x_n}}](\ell)$ is the inverse of the covariance matrix associated with the *suppressed* power spectra at multipole ℓ, and the difference between suppressed and unsuppressed power spectra is denoted by $\Delta P = P_{\mathrm{sup}} - P_{\mathrm{unsup}}$. In Eq. 5.10, f_{sky} denotes the fraction of the sky covered by the survey. The quantity $\sqrt{\Delta\chi^2}$ then represents the observable difference between the suppressed and unsuppressed models in units of the statistical uncertainty of the survey. That is, if the observed power spectrum matches the unsuppressed (ΛCDM) power spectrum then the suppressed model could be ruled out at a significance of $\sqrt{\Delta\chi^2}$ sigmas.

Second, we employ the Fisher matrix to estimate the statistical constraints that a future survey will be able to place on the cutoff parameter k_c. The Fisher matrix is defined as

$$F_{\alpha\beta} = \sum_{\ell=2}^{\ell_{\mathrm{max}}}(2\ell+1)f_{\mathrm{sky}}\sum_{\mathrm{i,j,m,n}}\frac{\partial P_{\mathrm{x_i x_j}}}{\partial p_\alpha}\mathbf{C}^{-1}[P_{\mathrm{x_i x_j}}, P_{\mathrm{x_m x_n}}]\frac{\partial P_{\mathrm{x_m x_n}}}{\partial p_\beta} + F^{\mathrm{P}}_{\alpha\beta}. \tag{5.11}$$

The parameters of the model are p_α and p_β, with Greek indices labeling the model parameters. For all our observables we set $\ell_{\mathrm{max}} = 300$; this ensures that the assumptions of weak lensing and Gaussian statistics are valid [32; 44; 165; 185; 192] and that modeling the galaxy distribution with a simple linear bias is appropriate; most interesting models of power suppression are constrained by multipoles $\ell < 30$, firmly in the linear galaxy bias regime, so we expect that our scale-independent bias assumption is well-founded.

The inverse of the Fisher matrix is an estimate of the parameter covariance near the maximum of the likelihood, i.e. at the fiducial values of the parameters. The measurement error on parameter α marginalized over all other parameters is

$$\sigma(p_\alpha) = [F^{-1}]_{\alpha\alpha}. \tag{5.12}$$

Throughout this work we assume flat spatial geometry and allow seven cosmological parameters to vary about the following fiducial values: $\Omega_{\rm m}h^2 = 0.13$, $w_0 = -1$, $w_{\rm a} = 0.0$, $\Omega_{\rm b}h^2 = 0.0223$, $n_{\rm s} = 0.96$, $\ln(\Delta^2_{\mathcal{R}}) = -19.953$, and $\Omega_\Lambda = 0.73$.

Gaussian priors on the parameters are incorporated into the Fisher analysis via the second term in Eq. 5.11. In all our forecasts we use the following priors: $\Delta\Omega_{\rm m}h^2 = 0.007$, $\Delta\Omega_{\rm b}h^2 = 0.001$, $\Delta n_{\rm s} = 0.04$, $\Delta\ln(\Delta^2_{\mathcal{R}}) = 0.1$; these are comparable to current marginalized constraints on these parameters [102] and so represent a conservative estimate for the statistical uncertainty on cosmology that will be achieved in advance of LSST or Euclid.

5.3 RESULTS

5.3.1 Effect of Suppression on Power Spectra

Figure 24 illustrates the effect of a particularly aggressive model of power suppression on the observables. The fractional change to $P^{\kappa\kappa}$ (dashed curves), P^{gg} (solid curves), and $P^{\kappa g}$ (dot-dashed curves) are each plotted as a function of multipole. In calculating the power spectra plotted in Fig. 24, we have used four tomographic bins, evenly spaced in the range $0 < z < 3$, for the distribution of both galaxy correlation sources and shear sources, so that the thin blue curves trace the change to each signal in the tomographic bin with redshift boundaries $0.75 < z < 1.5$, and the thick red curves trace changes in the bin with redshift boundaries $1.5 < z < 2.25$.

Two features of this figure are particularly worthy of note:

1. The signal at high redshift is more sensitive to large-scale primordial power suppression than the signal at low redshift, irrespective of the observable.
2. Galaxy-galaxy power spectra are more sensitive than shear-shear.

Both of these features are simple to understand. For a fixed angular scale, larger redshift corresponds to larger physical scale, so that for $i > j$, $P^{\rm x_i x_i}$ probes the matter distribution on larger scales than $P^{\rm x_j x_j}$ and thus $P^{\rm x_i x_i}$ will be more dramatically affected by large-scale power suppression.

Galaxy-galaxy power spectra are more sensitive to large-scale features in $\mathrm{P}_\delta(k)$ than shear-shear for a related reason. In Fig. 25 we compare the weight functions (defined by Eqns. 5.5 and 5.6) in each of four tomographic bins for $P^{\kappa\kappa}$ (dashed green) and P^{gg} (solid blue). Since W_i^g peaks at larger redshift than $\mathrm{W}_\mathrm{i}^\kappa$, P^{gg} is comparatively more sensitive to physics at high redshift than $P^{\kappa\kappa}$. This is simply because the galaxy-galaxy auto-power spectrum in a particular redshift bin probes the clustering properties of matter *within* that bin, whereas the shear-shear auto-power spectrum probes the distribution of matter that lies *in between* the source galaxies and the telescope. Thus, as discussed above, for a fixed angular scale P^{gg} is comparatively more sensitive to gravitational clustering on large scales than $P^{\kappa\kappa}$, and therefore large-scale power suppression will induce a greater fractional change in P^{gg} than $P^{\kappa\kappa}$.

5.3.2 Detectability

In this section we employ the $\Delta\chi^2$ technique described in §5.2.4 to study the ability of a very-wide-area photometric survey to distinguish between power spectra in standard ΛCDM and a model in which primordial power is suppressed on scales $k \lesssim k_c$ according to Eq. 5.1. Our chief result for this technique is plotted in Figure 26. The square root of $\Delta\chi^2$, computed via Eq. 5.10, appears on the vertical axis; the log of the cutoff scale appears on the horizontal axis. We have used the suppressed-model power spectra to calculate the covariance matrix in Eq. 5.10, so that Fig. 26 is suited to answer the following question: if the true model of the primordial power spectrum is unsuppressed, so that a LSST- or Euclid-like survey observes the large scale galaxy clustering statistics and cosmic shear signal predicted by standard ΛCDM, then as a result of such observations to what confidence could we rule out a given model of power suppression (i.e. a given value of k_c)? Thus the vertical axis values in Fig. 26 represent the confidence (the "number of sigmas") with which the suppressed model can be ruled out by planned imaging surveys. With the dotted magenta line we have plotted the detectability of power suppression with a spectroscopic survey similar to BigBOSS, corresponding to $f_{\mathrm{sky}} = 0.5$ and $N_\mathrm{A} = 0.5$ gal/arcmin^2; we have used the same galaxy distribution for our BigBOSS calculations as the $n(z)$ we used for our imaging survey

to facilitate a direct comparison between these two surveys. The technique used for the calculation of $\Delta\chi^2$ for the case of a spectroscopic survey is straightforward; we refer the reader to Ref. [57] for details. The detectability of power suppression with an LSST- or Euclid-like survey using $P^{\kappa\kappa}$ is shown by the dashed green curve, using P^{gg} the dot-dashed blue curve, and a joint analysis the solid red curve; detectability levels of 1σ and 3σ are delineated by solid, gray horizontal lines. There is little added advantage that a joint analysis has over using the galaxy-galaxy power spectra alone because P^{gg} is much more sensitive to horizon-size scales than $P^{\kappa\kappa}$ (see the discussion in §5.2.4).

Of particular interest here is the comparison between the dotted and solid curves, which can be thought of as contrasting the constraining power of the most ambitious large-scale spectroscopic and photometric surveys that will be undertaken over the next ten years. In this context, the advantage a spectroscopic redshift survey has over a photometric imaging survey is that the redshift survey has many more radial modes available to probe very large scales: photometric redshift uncertainty restricts the sampling of the radial signal to a small handful of tomographic redshift bins. Thus in general, the constraining power of an imaging survey increases with the number of tomographic bins as finer binning allows for more detailed study of the redshift evolution of the signal. This information eventually saturates and further refinement of the binning ceases to significantly improve the constraints. We find very little improvement in the $\Delta\chi^2$ results beyond $N_s = 4$ tomographic bins used for cosmic shear and $N_g = 8$ redshift bins for galaxy correlations. The limitation to the radial information that necessarily comes from using photometric data is more than compensated for by the greater surface density of sources that will be observed by LSST or Euclid. This may seem surprising since power suppression primarily affects multipoles $\ell \lesssim 30$ where cosmic variance is typically thought to dominate the errors. However, because most of the constraining power on primordial power suppression comes from sources at high redshift where the surface density of galaxies is quite sparse, shot noise is significant and imaging surveys will be able to exploit this relative advantage to distinguish at the 3σ level between ΛCDM and models of power suppression that are favored by the $S_{1/2}$ statistic, $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) \gtrsim -2.7$. Models that are mutually favored by both the CMB C_ℓ's and $S_{1/2}$, $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) \lesssim -3.3$, will not produce an effect that will be statistically significant in data sets that will be obtained

by LSST, Euclid, or BigBOSS, and hence these models will remain inaccessible to the galaxy surveys currently planned to take place within the next decade.

5.3.3 Statistical Constraints on k_c

In this section we present our results for the second method we used to assess the detectability of primordial power suppression, in which we employ the Fisher matrix formalism to forecast statistical constraints on the cutoff parameter k_c. The advantage this approach has over the method described in §5.3.2 is that the Fisher formalism provides a natural way to account for degeneracies between k_c, cosmological parameters, and galaxy bias, as well as a way to estimate the significance of photo-z uncertainty. We use seven cosmological parameters in our Fisher analysis with the same fiducial values and priors specified in §5.2.4.

The results from this calculation appear in Figure 27. On the horizontal axis is the log of the cutoff scale, on the vertical axis the statistical constraint on $\log_{10}(k_c)$. Constraints on $\log_{10}(k_c)$ when only convergence power spectra are used appear as the dashed green curve, using only P^{gg} as the dot-dashed blue curve, and a joint analysis as the solid red curve. Just as we found for our $\Delta\chi^2$ results in §5.3.2, the constraining power on the cutoff saturates at $N_g = 8$ tomographic redshift bins for galaxy power spectra and $N_s = 4$ redshift bins for cosmic shear.

The results presented in Fig. 27 are suited to answer the following question: if the true primordial power spectrum is, in fact, exponentially suppressed beyond some scale k_c, so that a LSST- or Euclid-like survey observes the large scale galaxy clustering statistics and cosmic shear signal predicted by the suppressed model, then to what statistical precision could the parameter k_c be constrained by such an observation? Thus this calculation is complementary to the one presented in the previous section in the following sense: results in §5.3.2 pertain to the difference between a given power suppression model and ΛCDM, whereas results in this section pertain to the difference between one power suppression model and another (nearby in k_c parameter space) power suppression model. The salient conclusions that can be drawn from Fig. 27 are similar to those in Fig. 26; when $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) = -2.7$ a joint analysis provides a relatively tight 7% constraint on the cutoff parameter; for power suppression

models that are mutually favored by both the $S_{1/2}$ statistic and the CMB multipoles the cutoff is on larger scales, $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) = -3.3$, where the constraining power of a survey such as LSST or Euclid is comparably weak.

In order to perform these calculations, the derivatives appearing in Eq. 5.11 must be evaluated numerically; we found that our results are robust to changes in both step-size as well as the choice to compute one- or two-sided derivatives for all parameters in the analysis. In particular, the statistical constraints on k_c are in principle asymmetric because constraining power varies monotonically with the scale of the cutoff. However, we find that the constraints vary sufficiently slowly with k_c to limit the effect of the asymmetry to a correction of only a few percent over the relevant range of parameter space.

Within reasonable levels, photometric redshift uncertainty turns out to have very little effect on the results in Fig. 27. We have checked this conclusion in two distinct ways. First, we parametrize photo-z uncertainty as described in Ref. [119]. Briefly, the redshift evolution of each of the functions σ_z and z^{bias} is modeled by linearly interpolating among a set of 31 control points, one at each interval of $\delta z = 0.1$ between $z = 0$ and $z = 3$, where the values at the control points are given by $\sigma_z = A(1+z)$ and $z^{\mathrm{bias}} = 0$, with $A = 0.05$ as our standard value of the photo-z spread at redshift zero. These $31 \times 2 = 62$ control points then serve as photo-z uncertainty parameters in the Fisher analysis. We find that the power spectrum cutoff parameters exhibit very little degeneracy with photometric redshift parameters, so that the constraints are not degraded significantly by marginalizing over these or more complex photometric redshift parameterizations. Second, we studied the sensitivity of the constraints to the value of $A = \sigma_z/(1+z)$. The chief effect that varying A has on our results comes from restricting the statistically independent information in the tomography, i.e. larger values of A lead to constraining power that saturates at smaller numbers of tomographic redshift bins because photo-z uncertainty smears out correlations along the line of sight and thus restricts the number of radial modes available to probe cosmology. However, for reasonable values of A this effect is quite small: $\sigma(\log_{10}(k_c))$ only changes by roughly ten percent when A varies between the optimistic value of $A = 0.03$ and the quite pessimistic value of $A = 0.15$. As the effect of photo-z uncertainty on the dark energy constraints is much more profound [5; 13; 64; 82; 108; 119; 196; 203], the photo-z calibration effort leading up to future wide-

area imaging surveys will very likely achieve the precision required to reach the constraining power illustrated in Figs. 26-27.

5.4 CONCLUSIONS AND DISCUSSION

We have studied the sensitivity of an all-sky galaxy imaging survey such as LSST or Euclid to the suppression of primordial power on scales comparable to the horizon size at the time of recombination. The models of suppression we investigated are motivated by the observed deficit in the two-point correlation function of the CMB temperature at large angles. In particular, we focused on constraining the comoving cutoff scale k_c at which exponential suppression of $\Delta^2_\delta(k)$ sets in.

The cutoff scale favored by the $S_{1/2}$ statistic alone is ~ 700 Mpc. We find that a LSST- or Euclid-like survey will be able to distinguish this model from ΛCDM at the 3σ level, and that if we do in fact live in a universe in which the primordial power spectrum is exponentially suppressed on comoving scales larger than ~ 700 Mpc, then a joint analysis of shear and galaxy correlations could provide 7% constraints on the cutoff parameter. However, planned galaxy surveys will not be able to discriminate between ΛCDM and models in which the suppression sets in at cutoff scales ~ 2.8 Gpc, as favored jointly by the C_ℓ's and $S_{1/2}$. The chief reason for the relatively weak constraining power on models with a cutoff on these larger scales is simply that the redshift range covered by the next generation of galaxy surveys is not deep enough to probe matter-power-spectrum modes larger than a few Gpc. Thus it may be necessary to rely on future observations of the CMB, particularly the polarization signal and its cross-correlation with the temperature (as studied in Ref. [126]), to test power suppression models on these very large scales.

In our forecasts of the constraints on power suppression we have not taken into account possible systematic errors that complicate any observation of galaxy clustering or cosmic shear on large scales. Any calibration error that varies across the sky over a difference of $\sim 20^\circ$ can interfere with a measurement of the large-scale clustering signal and therefore contribute to the error budget of a galaxy survey. For example, an effective change to the

magnitude limit of a survey can be induced by Galactic extinction, which may vary with the line of sight through the Milky Way. The relative frequency of star-galaxy misclassifications may effectively vary across the sky due to the spatial dependence of stellar demographics in the Galaxy. We leave a detailed assessment of possible sources of systematic error as a task for future work.

Finally, we have compared the sensitivity to large-scale features in $\Delta^2_\delta(k)$ of a spectroscopic redshift survey with characteristics similar to that of BigBOSS and future wide-area imaging surveys such as LSST or Euclid. We find that while a BigBOSS-like survey has a greater number of statistically independent radial modes with which to probe large scales, planned imaging surveys will be more sensitive to horizon-scale physics because of the greater surface density of high-redshift sources they will observe.

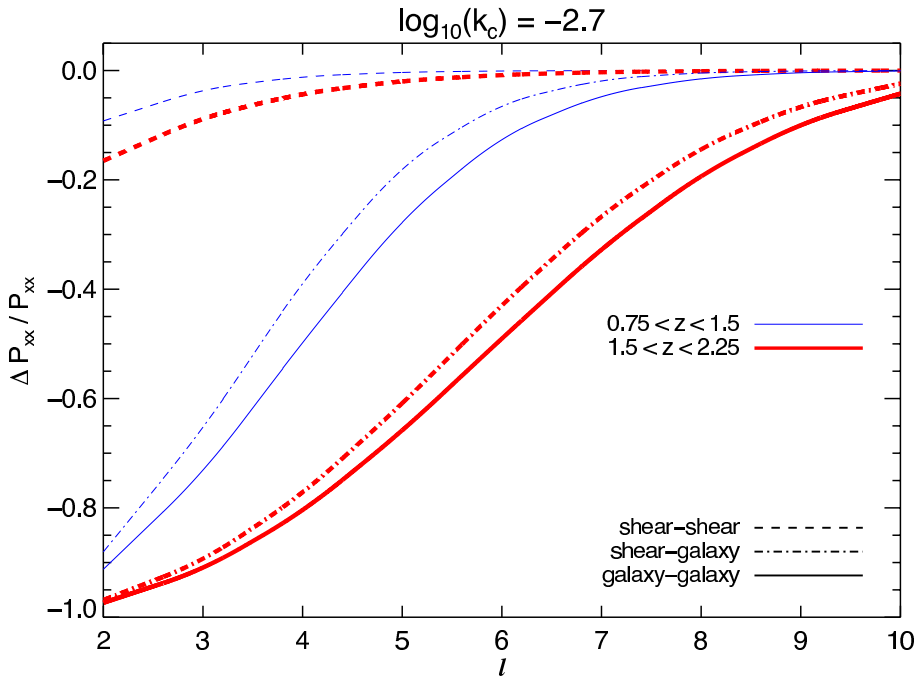

Figure 24: Toy demonstration of the effect of power suppression on the three sets of power spectra we study: galaxy-galaxy are plotted with solid curves, shear-shear with dashed, and shear-galaxy with dot-dashed. The fractional change to each observable, $(P^{x_i x_j}_{\rm sup} - P^{x_i x_j}_{\rm unsup})/P^{x_i x_j}_{\rm unsup}$, is plotted as a function of multipole. The fractional change to the signal at high redshift is plotted with thick, red lines for each observable; the signal at low redshift with thin, blue lines. High-redshift bins are more fractionally perturbed than low because for a fixed angular scale, larger redshift corresponds to larger physical scale. Galaxy-galaxy power correlations are more affected than shear-shear since the redshift kernel peaks at higher redshift for galaxy-galaxy (the lenses are in front of the galaxies).

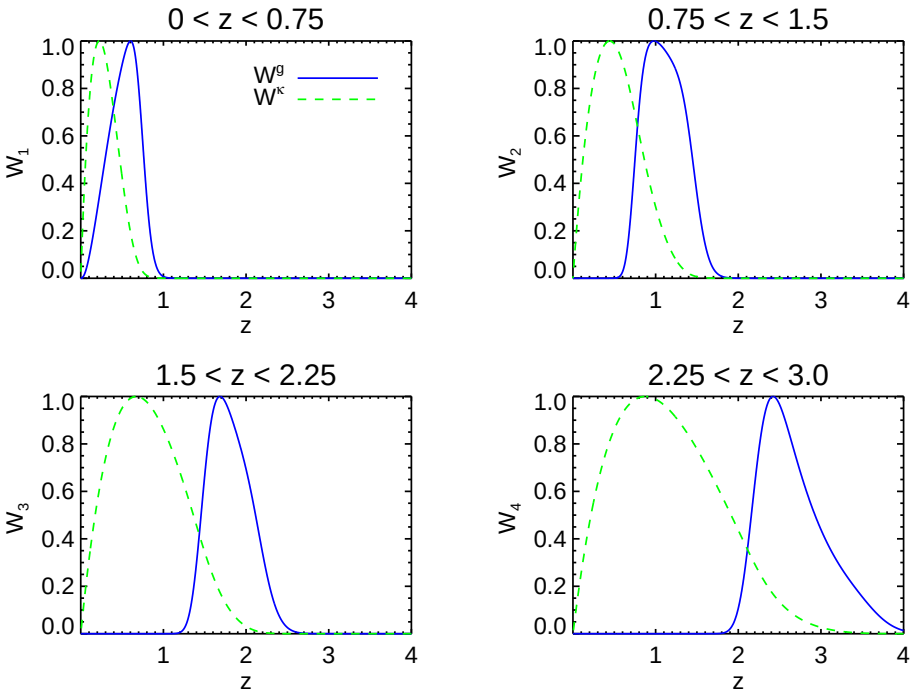

Figure 25: Comparison of the weight functions (Eqns. 5.5 & 5.6) for a toy set of galaxy-galaxy auto-power spectra (solid blue curves) and shear-shear auto-power spectra (dashed green curves), each divided into four tomographic bins. Each kernel has been normalized by its own maximum value to facilitate a direct comparison of the redshift evolution of the weight functions.

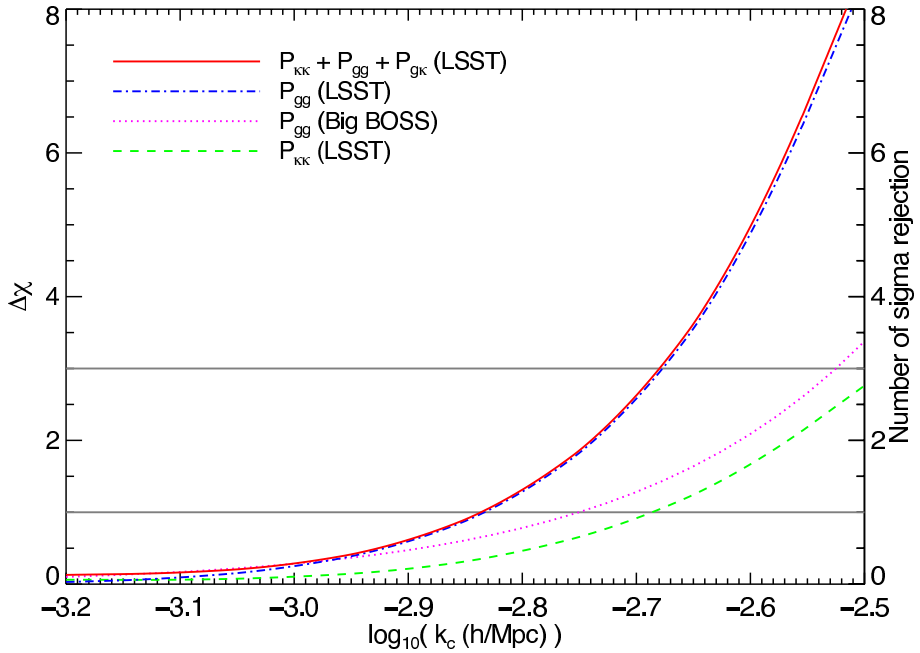

Figure 26: The cutoff scale at which primordial power suppression becomes significant appears on the horizontal axis. Along the vertical axis is $\Delta\chi = \sqrt{\Delta\chi^2}$, defined by Eq. 5.10, our first statistic quantifying the detectability of power suppression with future wide-area surveys. As discussed in § 5.3.1, galaxy clustering has significantly greater constraining power on primordial power suppression models than cosmic shear. Models of power suppression favored by the $S_{1/2}$ statistic alone, $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) = -2.7$, could be ruled out at the 3σ level by a joint analysis of an LSST- or Euclid-like survey; models with $\log_{10}(k_c/h\mathrm{Mpc}^{-1}) = -3.3$ that are mutually favored by the C_ℓ's and $S_{1/2}$ will be inaccessible to the surveys we consider.

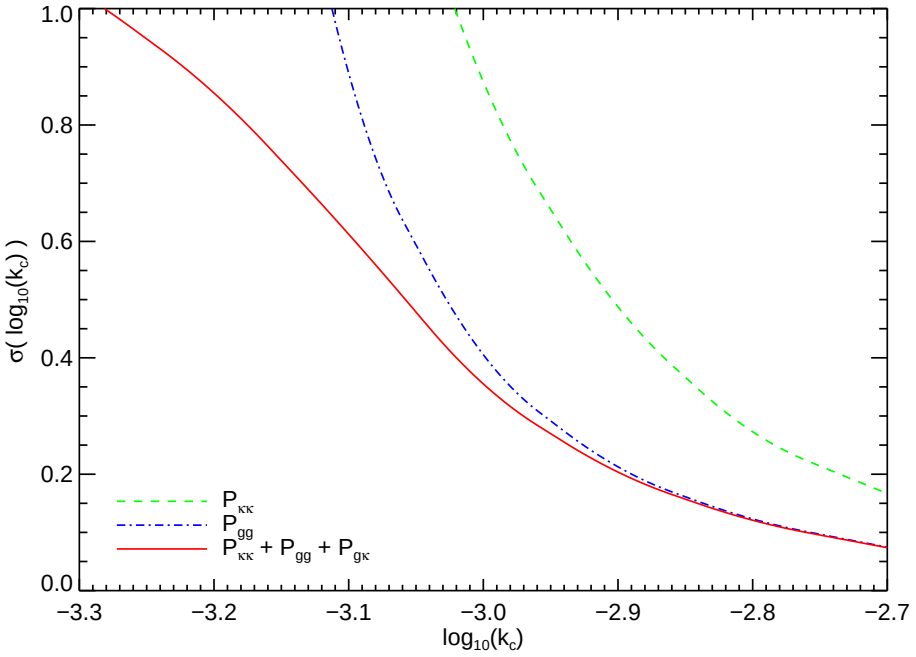

Figure 27: Plot of the statistical constraining power that future imaging surveys will have on the cutoff scale. We compute $\sigma(\log_{10}(k_c))$, the precision to which the scale of the cutoff can be measured, using the Fisher matrix formalism; forecasts for the constraints from an LSST- or Euclid-like imaging survey using galaxy clustering statistics alone appear as the dot-dashed blue curve, using cosmic shear alone as the dashed green curve, and with a joint analysis the solid red curve.

6.0 SUMMARY & CONCLUSIONS

In this dissertation I have explored the potential of the weak lensing signal that will be observed in planned and near future galaxy imaging surveys to yield new insight into ΛCDM. The purely geometric nature of weak lensing can be exploited to test general relativity on cosmological scales, a possibility I explored in chapter 2. Such tests are very promising, although their constraining power relies on being able to predict the matter power spectrum, $P_\delta(k)$, deep into the nonlinear regime where baryonic effects are significant. Characterizing the influence of baryons on $P_\delta(k)$ is challenging, even for state-of-the-art numerical simulations. However, I found that if baryonic influences are limited to changes in halo concentrations, then this source of uncertainty does not seriously threaten weak lensing tests of general relativity.

In chapter 4 I studied the related issue of the significance of matter power spectrum errors for the study of dark energy with weak lensing. In particular, I updated the $P_\delta(k)$ calibration requirements to the more realistic case of having simultaneous uncertainty in the photometric redshifts of the lensed galaxies. I found that the matter power spectrum will need to be predicted to a precision of 0.5% to wavenumbers $k \sim 5h\mathrm{Mpc}^{-1}$ in order for Stage IV experiments such as LSST to reach their target constraints on dark energy. As part of this investigation, I provided a blueprint for estimating both the precision and required scale for the $P_\delta(k)$ calibration as a function of the maximum multipole used in the lensing analysis (see Figures 21 and 22). This blueprint can be used by the community of numerical cosmologists studying this problem to set goalposts and assess their current achievements in terms of dark energy constraints.

A significant fraction of the literature in numerical cosmology will likely be dedicated to the problems presented in chapters 2 & 4 for years to come, reflecting the formidable

challenge presented by these calibration requirements. At the N-body level, the $P_\delta(k)$ calibration problem is well-defined and the brute force approach that has been adopted thus far (an aggressive campaign of gravity-only cosmological simulations geared towards providing a power spectrum emulator) continues to produce steady improvements in the achieved precision. With our current level of knowledge, it is more difficult to assess the future of the baryonic side of this program. For example, widespread agreement over the appropriate method to implement feedback mechanisms via sub-grid recipes has not yet been reached, and the sensitivity of baryonic effects to the underlying cosmology remains poorly tested. However, we find very promising the success of the self-calibration approach studied in chapters 2 & 4. While it remains to be seen whether other baryonic effects beyond halo concentration modification must be taken into account on scales of interest, halo concentration is already known to encode a significant fraction of the baryonic contribution to $P_\delta(k)$, and we have shown that such modifications do not present a serious threat to cosmic shear studies of dark energy or general relativity.

In chapter 5 I investigate the possibility of using large-scale structure measurements with LSST to detect deviations from the standard inflationary prediction for the primordial power spectrum. If exponentially suppressed primordial power is the reason behind the lack of correlation in the CMB temperature anisotropy on large angles, then signatures of this suppression should in principle manifest in the galaxy distribution and the cosmic shear signal observed by LSST. My results in chapter 5 demonstrate that this is indeed the case. However, I found that in order for weak lensing and galaxy correlations to provide precise constraints on the power suppression model, the suppression must be operative on scales sufficiently small to alter the acoustic peak structure of the CMB. It may be necessary to await the arrival of high redshift 21-cm power spectra to probe models of primordial power suppression on scales large enough so that the model under consideration accounts for the large-angle anomaly without sacrificing the fit to the acoustic peaks, although this approach remains to be tested.

In chapter 3 I provided a detailed study of the influence of photometric redshift errors on the cosmic shear signal. I found that uncalibrated outlier populations in the photo-z distribution will need to be kept at or below per-galaxy rates of 0.1% in order to ensure that

they do not contribute significantly to the dark energy error budget. Excising galaxies at high- and low-redshifts, where attaining photo-z precision is the most challenging, comes at only a modest cost to the statistical constraints. This strategy may alleviate the challenging calibration demands presented in chapter 3, but a careful examination of Figure 11 shows that even with this excision, contributions from outlier populations at the sub-percent level still induce dark energy systematics at threatening levels.

To date, the most common treatment in the literature of the full distribution of photometric redshifts divides the galaxy population into a core sample and an outlier sample, quantifies the outlier rate and focuses on characterizing the properties of the core sample. Outlier rates in contemporary studies persist at the level of a few percent, and these rates have shown no sign of decreasing by an order of magnitude for the foreseeable future. Thus the implications of the results presented in chapter 3 are that a more sophisticated photo-z error analysis will ultimately be required to keep dark energy systematics due to photo-z outliers from dominating the error budget. If outlier rates persist at contemporary levels, it will be necessary to incorporate these populations into the core distribution and accurately characterize their properties. The emerging cross-correlation technique appears to be one of the most promising new methods to make such a characterization possible. A thorough examination of the ability to use this technique for outlier calibration is a specific direction for future work in this field that may provide the leap forward that will be necessary to solve this challenging problem.

APPENDIX

THE NONLINEAR POWER SPECTRUM AND PHOTOMETRIC REDSHIFT CALIBRATION REQUIREMENTS

Much of the constraining power of weak lensing surveys arises from measurements on scales where the structures causing the deflections are undergoing nonlinear gravitational evolution [e.g. 80]. Restricting consideration to large scales significantly degrades cosmological constraints [e.g. 63; 78; 80; 159; 197], so it is necessary to model nonlinear evolution in order to utilize weak lensing to constrain dark energy. At least three approximate and related techniques are in common use: (1) the fitting formula of Peacock and Dodds [141], which is based on the HKLM method [61]; (2) the halo model [33; 117; 157; 162; 164]; and (3) the fitting formula of Smith et al. [171]. The works of Ma et al. [119] and Ma and Bernstein [118] specifying requirements for photometric redshift calibration employ the Peacock and Dodds [141] relation.

In the course of our study, we have recomputed the photometric redshift calibration requirements using each of the three approximate techniques mentioned in the previous paragraph. In the limit of perfect knowledge of the photometric redshift distribution, each of these fitting formulas gives nearly identical dark energy constraints. However, we have found that the photometric redshift calibration requirements have a strong dependence upon the method used to model nonlinear structure. We summarize this finding in Figure 28 where we display contours of constant degradation in the statistical error on w_a as a function of both the prior on the bias $\Delta z^{\rm bias}$ and the prior on the dispersion $\sigma_{\rm z}$. In other words, we show contours of $\sigma(w_{\rm a})$ in units of the statistical constraint on $w_{\rm a}$ in the limit that the photo-z

distribution parameters are known perfectly prior to the weak lensing analysis, $\sigma_{\rm perf}(w_{\rm a})$. We assume that the same priors are applied to all of our 31 dispersion parameters and 31 bias parameters at each redshift bin (see § 3.2.1.1). We summarize our findings in this way so that these results can be compared directly to Figure 7 in Ma et al. [119]. To make the comparison as direct as possible, we have computed these forecasts using the fiducial cosmology and experimental setup of Ma et al. [119], which differs slightly from those considered in the main text. In this appendix only, our fiducial cosmology is $\omega_{\rm m} = 0.14$, $\omega_{\rm b} = 0.024$, $n_{\rm s} = 1.0$, $\Delta^2_{\mathcal{R}} = 2.4 \times 10^{-5}$ (giving $\sigma_8 \simeq 0.91$), and $\Omega_{\rm DE} = 0.73$ combined with experimental parameters of $f_{\rm sky} = 0.1$ and $N^A = 55$ arcmin^{-2}.

The upper left panel in Figure 28 shows photo-z calibration requirements estimated using the Peacock and Dodds [141] treatment of nonlinear power. This panel shows nearly identical results to those in Figure 7 of [119] so that this panel validates our methods and provides a useful baseline to compare with the other panels. According to this result, ensuring that constraints on $w_{\rm a}$ are not degraded by more than a factor of two requires knowing the photo-z dispersion and bias to roughly $\sim 1\%$ prior to undertaking the weak lensing analysis. The upper, right panel of Fig. 28 shows the same requirements constructed using the halo model for nonlinear clustering. In the limit of restrictive prior knowledge of the photo-z distribution the Peacock and Dodds [141] and halo model results yield nearly the same constraints. When the photo-z distributions have relatively unrestrictive priors, the two techniques yield moderately different levels of projected degradation with, for example, uncertainty in the photo-z dispersion of $\Delta(\sigma_{\rm z}) \approx 1$ corresponding to a factor of ten degradation in the Peacock and Dodds [141] case but a factor of six degradation in the halo model calculation.

The largest differences among the forecasts comes from comparing the requirements using Peacock and Dodds [141] to those computed using the Smith et al. [171] fit. As with the halo model comparison, the different techniques agree well when prior knowledge of the photo-z distribution is very restrictive; thus as long as degradations due to photo-z uncertainty are $\lesssim 10 - 20\%$ then it does not matter which technique one uses to predict the nonlinear evolution. It is interesting that the constraints in the case of the halo model treatment degrade significantly less rapidly as prior knowledge becomes less and less restrictive. Turning to the Smith et al. [171] fit, one would conclude that ensuring less than a factor of two

degradation on the $w_{\rm a}$ constraint requires $\sim 18\%$ knowledge of the dispersion and $\sim 40\%$ knowledge of the bias as compared to the $\sim 1\%$ requirements that result from the Peacock and Dodds [141] analysis.

Clearly, at most one of these treatments can represent the growth of cosmic structure faithfully. In the main text, we presented results using the Smith et al. [171] formula because these authors perform a detailed numerical study that finds the Peacock and Dodds [141] and simple implementations of the halo model to be imprecise on scales relevant for cosmic shear cosmology. In the context of these fitting formulae, we find that Smith et al. [171] predicts greater power than Peacock and Dodds [141] on scales most relevant to lensing ($0.1 \lesssim k/h{\rm Mpc}^{-1} \lesssim 10$), particularly at high redshift. At this point, it is not possible to make a firm statement as to which approach is correct, but an exhaustive simulation program similar to that being carried out by Heitmann et al. [66, 67, 69] may be capable of providing a more definitive resolution in the case of dissipationless evolution. Additional effort will be needed to treat any modifications induced by the baryonic component of the universe [59; 89; 155; 176; 191; 197; 199].

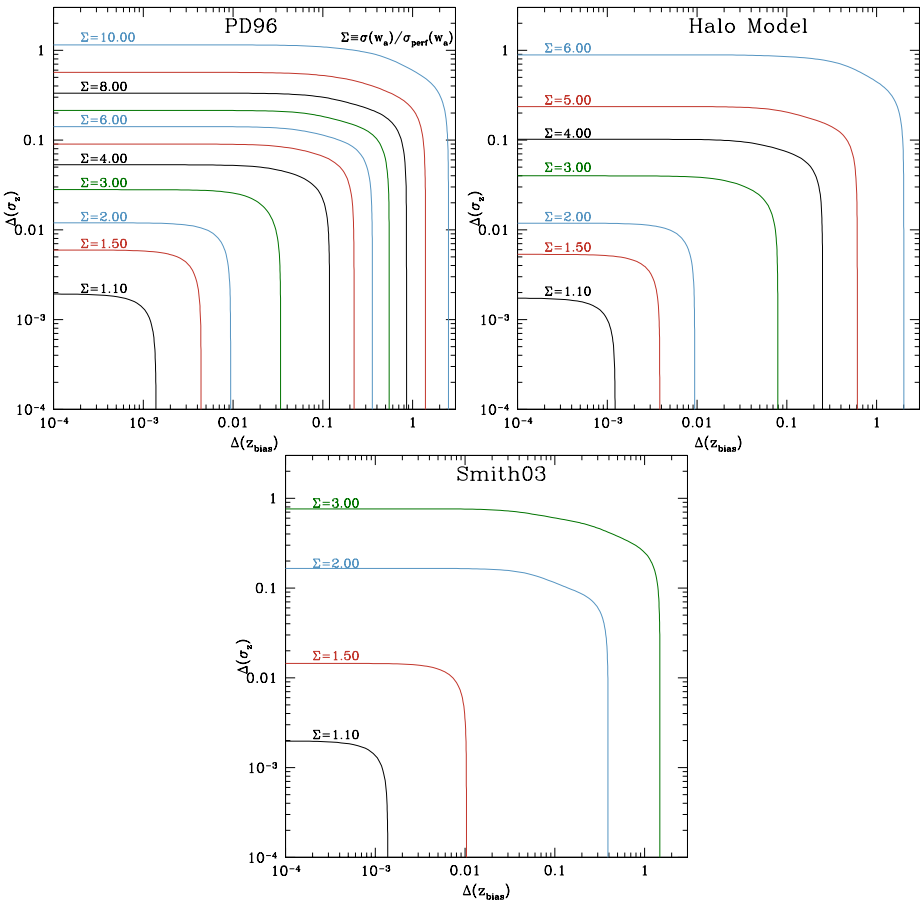

Figure 28: Contour plots for the level of w_a constraint degradation as a function of priors on the photometric redshift scatter σ_z and bias $z^{\rm bias}$. In this case, the priors are applied uniformly to the photometric redshift parameters are each redshift. The contours demarcate equal parameter degradation defined as the error on w_a after marginalizing over photometric reshift uncertainties. We show constraints in units of the equivalant constraint in the limit of perfect knowledge of photometric redshift parameters, $\Sigma \equiv \sigma(w_a)/\sigma_{\rm perf}(w_a)$. The *upper, left* panel was computed using the Peacock and Dodds [141] fitting formula for the nonlinear power spectrum of density fluctuations and amounts to a near reproduction of the right panel of Figure 7 in Ma et al. [119]. The *upper, right* panel was computed using the halo model as described in Zentner et al. [197]. The *bottom* panel was computed using the Smith et al. [171] relation for the nonlinear power spectrum of density fluctuations. Significant differences between the levels of degradation are evident. Note that in this figure, we use a different set of cosmological and experimental parameters so that this result is directly comparable to those in Figure 7 of Ma et al. [119].

BIBLIOGRAPHY

[1] K. Abazajian, E. R. Switzer, S. Dodelson, K. Heitmann, and S. Habib. Nonlinear cosmological matter power spectrum with massive neutrinos: The halo model. *Phys. Rev. D*, 71(4):043507, February 2005. doi: 10.1103/PhysRevD.71.043507.

[2] F. B. Abdalla, M. Banerji, O. Lahav, and V. Rashkov. A Comparison of Six Photometric Redshift Methods Applied to 1.5 Million Luminous Red Galaxies. *ArXiv e-prints*, December 2008.

[3] P Abell et al. LSST Science Book, Version 2.0. *[arXiv:0912.0201]*, December 2009.

[4] A. Albrecht, G. Bernstein, R. Cahn, W. L. Freedman, J. Hewitt, W. Hu, J. Huth, M. Kamionkowski, E. W. Kolb, L. Knox, J. C. Mather, S. Staggs, and N. B. Suntzeff. Report of the Dark Energy Task Force. *(astro-ph/0609591)*, September 2006.

[5] A. Albrecht, G. Bernstein, R. Cahn, W. L. Freedman, J. Hewitt, W. Hu, J. Huth, M. Kamionkowski, E. W. Kolb, L. Knox, J. C. Mather, S. Staggs, and N. B. Suntzeff. Report of the Dark Energy Task Force. *ArXiv Astrophysics e-prints*, September 2006.

[6] P. Astier, J. Guy, N. Regnault, and et al. The Supernova Legacy Survey: measurement of $\Omega_{\rm M}$, Ω_Λ and w from the first year data set. *Astron. Astrophys.*, 447:31–48, February 2006. doi: 10.1051/0004-6361:20054185.

[7] M. L. Balogh, F. R. Pearce, R. G. Bower, and S. T. Kay. Revisiting the cosmic cooling crisis. *Mon. Not. R. Astron. Soc.*, 326:1228–1234, October 2001. doi: 10.1046/j.1365-8711.2001.04667.x.

[8] M. Banerji, F. B. Abdalla, O. Lahav, and H. Lin. Photometric redshifts for the Dark Energy Survey and VISTA and implications for large-scale structure. *MNRAS*, 386: 1219–1233, May 2008. doi: 10.1111/j.1365-2966.2008.13095.x.

[9] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner. Spontaneous creation of almost scale-free density perturbations in an inflationary universe. *Phys. Rev. D*, 28:679–693, August 1983.

[10] D. Baumann and A. Cooray. CMB-induced cluster polarization as a cosmological

probe. *New Astronomy Reviews*, 47:839–843, November 2003. doi: 10.1016/j.newar.2003.07.013.

[11] J. Benjamin, C. Heymans, E. Semboloni, L. van Waerbeke, H. Hoekstra, T. Erben, M. D. Gladders, M. Hetterscheidt, Y. Mellier, and H. K. C. Yee. Cosmological constraints from the 100-square degree weak-lensing survey. *Mon. Not. R. Astron. Soc.*, 381:702–712, October 2007. doi: 10.1111/j.1365-2966.2007.12202.x.

[12] C. L. Bennett, R. S. Hill, G. Hinshaw, D. Larson, K. M. Smith, J. Dunkley, B. Gold, M. Halpern, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, S. S. Meyer, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Are There Cosmic Microwave Background Anomalies? *Astrophys. J.*, 192:17–+, February 2011. doi: 10.1088/0067-0049/192/2/17.

[13] G. Bernstein and D. Huterer. Catastrophic photometric redshift errors: weak lensing survey requirements. *Mon. Not. R. Astron. Soc.*, 401:1399, February 2010.

[14] Gary M. Bernstein and B. Jain. Dark energy constraints from weak lensing cross-correlation cosmography. *Astrophys. J.*, 600:17–25, 2004.

[15] S. Bhattacharya, K. Heitmann, M. White, Z. Lukić, C. Wagner, and S. Habib. Mass Function Predictions Beyond ΛCDM. *ApJ*, 732:122, May 2011. doi: 10.1088/0004-637X/732/2/122.

[16] S. Bird, M. Viel, and M. G. Haehnelt. Massive Neutrinos and the Non-linear Matter Power Spectrum. *ArXiv e-prints*, September 2011.

[17] M. Bolzonella, J.-M. Miralles, and R. Pelló. Photometric redshifts based on standard SED fitting procedures. *AAP*, 363:476–492, November 2000.

[18] D. Boyanovsky, H. J. de Vega, and N. G. Sanchez. CMB quadrupole suppression. II. The early fast roll stage. *Phys. Rev. D*, 74(12):123007–+, December 2006. doi: 10.1103/PhysRevD.74.123007.

[19] G. B. Brammer, P. G. van Dokkum, and P. Coppi. EAZY: A Fast, Public Photometric Redshift Code. *ApJ*, 686:1503–1513, October 2008. doi: 10.1086/591786.

[20] S. L. Bridle, A. M. Lewis, J. Weller, and G. Efstathiou. Reconstructing the primordial power spectrum. *Mon. Not. R. Astron. Soc.*, 342:L72–L78, July 2003. doi: 10.1046/j.1365-8711.2003.06807.x.

[21] J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. *Mon. Not. R. Astron. Soc.*, 321:559–575, March 2001.

[22] E. F. Bunn. Probing the Universe on gigaparsec scales with remote cosmic microwave background quadrupole measurements. *Phys. Rev. D*, 73(12):123517, June 2006. doi: 10.1103/PhysRevD.73.123517.

[23] E. F. Bunn and A. Bourdon. Contamination cannot explain the lack of large-scale power in the cosmic microwave background radiation. *Phys. Rev. D*, 78(12):123509–+, December 2008. doi: 10.1103/PhysRevD.78.123509.

[24] D. A. Buote, F. Gastaldello, P. J. Humphrey, L. Zappacosta, J. S. Bullock, F. Brighenti, and W. G. Mathews. The X-Ray Concentration-Virial Mass Relation. *(astro-ph/0610135)*, October 2006.

[25] L. Casarini, G. La Vacca, L. Amendola, S. A. Bonometto, and A. V. Macciò. Non-linear weak lensing forecasts. *J. Cosmol. Astropart. Phys.*, 3:26–+, March 2011. doi: 10.1088/1475-7516/2011/03/026.

[26] L. Casarini, A. V. Macciò, S. A. Bonometto, and G. S. Stinson. High-accuracy power spectra including baryonic physics in dynamical Dark Energy models. *Mon. Not. R. Astron. Soc.*, 412:911–920, April 2011. doi: 10.1111/j.1365-2966.2010.17948.x.

[27] M. Chevallier and D. Polarski. Accelerating Universes with Scaling Dark Matter. *International Journal of Modern Physics D*, 10:213–223, 2001. doi: 10.1142/S0218271801000822.

[28] A. L. Coil, M. Davis, D. S. Madgwick, J. A. Newman, C. J. Conselice, M. Cooper, R. S. Ellis, S. M. Faber, D. P. Finkbeiner, P. Guhathakurta, N. Kaiser, D. C. Koo, A. C. Phillips, C. C. Steidel, B. J. Weiner, C. N. A. Willmer, and R. Yan. The DEEP2 Galaxy Redshift Survey: Clustering of Galaxies in Early Data. *ApJ*, 609:525–538, July 2004. doi: 10.1086/421337.

[29] A. L. Coil, J. A. Newman, M. C. Cooper, M. Davis, S. M. Faber, D. C. Koo, and C. N. A. Willmer. The DEEP2 Galaxy Redshift Survey: Clustering of Galaxies as a Function of Luminosity at z = 1. *ApJ*, 644:671–677, June 2006. doi: 10.1086/503601.

[30] A. A. Collister and O. Lahav. ANNz: Estimating Photometric Redshifts Using Artificial Neural Networks. *PASPs*, 116:345–351, April 2004. doi: 10.1086/383254.

[31] C. R. Contaldi, M. Peloso, L. Kofman, and A. Linde. Suppressing the lower multipoles in the CMB anisotropies. *Journal of Cosmology and Astro-Particle Physics*, 7:2–+, July 2003. doi: 10.1088/1475-7516/2003/07/002.

[32] A. Cooray and W. Hu. Power Spectrum Covariance of Weak Gravitational Lensing. *Astrophys. J.*, 554:56–66, June 2001. doi: 10.1086/321376.

[33] A. Cooray and R. Sheth. Halo models of large scale structure. *Phys. Rep.*, 372:1–129, December 2002.

[34] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman. Uncorrelated universe: Statistical anisotropy and the vanishing angular correlation function in WMAP years 1 3. *Phys. Rev. D*, 75(2):023507–+, January 2007. doi: 10.1103/PhysRevD.75.023507.

[35] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman. No large-angle correlations on the non-Galactic microwave sky. *Mon. Not. R. Astron. Soc.*, 399:295–303, October 2009. doi: 10.1111/j.1365-2966.2009.15270.x.

[36] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman. Large-Angle Anomalies in the CMB. *Advances in Astronomy*, 2010:78–+, 2010. doi: 10.1155/2010/847541.

[37] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman. Bias in low-multipole CMB reconstructions. *ArXiv e-prints*, March 2011.

[38] J. Coupon, O. Ilbert, M. Kilbinger, H. J. McCracken, Y. Mellier, S. Arnouts, E. Bertin, P. Hudelot, M. Schultheis, O. Le Fèvre, V. Le Brun, L. Guzzo, S. Bardelli, E. Zucca, M. Bolzonella, B. Garilli, G. Zamorani, A. Zanichelli, L. Tresse, and H. Aussel. Photometric redshifts for the CFHTLS T0004 deep and wide fields. *AAP*, 500:981–998, June 2009. doi: 10.1051/0004-6361/200811413.

[39] C. E. Cunha, M. Lima, H. Oyaizu, J. Frieman, and H. Lin. Estimating the redshift distribution of photometric galaxy samples - II. Applications and tests of a new method. *MNRAS*, 396:2379–2398, July 2009. doi: 10.1111/j.1365-2966.2009.14908.x.

[40] C. E. Cunha, D. Huterer, M. T. Busha, and R. H. Wechsler. Sample variance in photometric redshift calibration: cosmological biases and survey requirements. *ArXiv e-prints*, September 2011.

[41] T. Dahlen, B. Mobasher, S. Jouvel, J.-P. Kneib, O. Ilbert, S. Arnouts, G. Bernstein, and J. Rhodes. Supernova Acceleration Probe: Investigating Photometric Redshift Optimization. *AJ*, 136:1361–1371, September 2008. doi: 10.1088/0004-6256/136/3/1361.

[42] S. Dodelson and A. Vallinotto. Learning from the scatter in type Ia supernovae. *Phys. Rev. D*, 74(6):063515–+, September 2006. doi: 10.1103/PhysRevD.74.063515.

[43] S. Dodelson and P. Zhang. Weak lensing bispectrum. *Phys. Rev. D*, 72(8):083001, October 2005. doi: 10.1103/PhysRevD.72.083001.

[44] S. Dodelson, C. Shapiro, and M. White. Reduced shear power spectrum. *Phys. Rev. D*, 73(2):023009–+, January 2006. doi: 10.1103/PhysRevD.73.023009.

[45] D. Dolney, B. Jain, and M. Takada. Effects of halo substructure on the power spectrum and bispectrum. *Mon. Not. R. Astron. Soc.*, 352:1019–1027, August 2004. doi: 10.1111/j.1365-2966.2004.07999.x.

[46] O. Doré, M. Martig, Y. Mellier, M. Kilbinger, J. Benjamin, L. Fu, H. Hoekstra, M. Schultheis, E. Semboloni, and I. Tereno. Testing Gravity with the CFHTLS-Wide Cosmic Shear Survey and SDSS LRGs. *ArXiv:0712.1599*, December 2007.

[47] G. Dvali, G. Gabadadze, and M. Porrati. 4D gravity on a brane in 5D Minkowski space. *Physics Letters B*, 485:208–214, July 2000.

[48] G. Efstathiou, Y.-Z. Ma, and D. Hanson. Large-angle correlations in the cosmic microwave background. *Mon. Not. R. Astron. Soc.*, 407:2530–2542, October 2010. doi: 10.1111/j.1365-2966.2010.17081.x.

[49] T. Eifler. Weak-lensing statistics from the Coyote Universe. *MNRAS*, pages 1490–+, September 2011. doi: 10.1111/j.1365-2966.2011.19502.x.

[50] D. J. Eisenstein, I. Zehavi, D. W. Hogg, and et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, November 2005. doi: 10.1086/466512.

[51] W. Fang, S. Wang, W. Hu, Z. Haiman, L. Hui, and M. May. Challenges to the DGP Model from Horizon-Scale Growth and Geometry. *ArXiv e-prints*, August 2008.

[52] T. Faulkner, M. Tegmark, E. F. Bunn, and Y. Mao. Constraining f(R) gravity as a scalar-tensor theory. *Phys. Rev. D*, 76(6):063505–+, September 2007. doi: 10.1103/PhysRevD.76.063505.

[53] S. M. Feeney, H. V. Peiris, and A. Pontzen. Avoiding bias in reconstructing the largest observable scales from partial-sky data. *ArXiv e-prints*, July 2011.

[54] R. Feldmann, C. M. Carollo, C. Porciani, S. J. Lilly, P. Capak, Y. Taniguchi, O. Le Fèvre, A. Renzini, N. Scoville, M. Ajiki, H. Aussel, T. Contini, H. McCracken, B. Mobasher, T. Murayama, D. Sanders, S. Sasaki, C. Scarlata, M. Scodeggio, Y. Shioya, J. Silverman, M. Takahashi, D. Thompson, and G. Zamorani. The Zurich Extragalactic Bayesian Redshift Analyzer and its first application: COSMOS. *MNRAS*, 372:565–577, October 2006. doi: 10.1111/j.1365-2966.2006.10930.x.

[55] B. Feng and X. Zhang. Double inflation and the low CMB quadrupole. *Physics Letters B*, 570:145–150, September 2003. doi: 10.1016/j.physletb.2003.07.065.

[56] L. Fu, E. Semboloni, H. Hoekstra, M. Kilbinger, L. van Waerbeke, I. Tereno, Y. Mellier, C. Heymans, J. Coupon, K. Benabed, J. Benjamin, E. Bertin, O. Doré, M. J. Hudson, O. Ilbert, R. Maoli, C. Marmo, H. J. McCracken, and B. Ménard. Very weak lensing in the CFHTLS wide: cosmology from cosmic shear in the linear regime. *Astron. Astrophys.*, 479:9–25, February 2008. doi: 10.1051/0004-6361:20078522.

[57] C. Gibelyou, D. Huterer, and W. Fang. Detectability of large-scale power suppression in the galaxy distribution. *Phys. Rev. D*, 82(12):123009–+, December 2010. doi: 10.1103/PhysRevD.82.123009.

[58] T. Guillet, R. Teyssier, and S. Colombi. The effect of baryons on the variance and the skewness of the mass distribution in the Universe at small scales. *Mon. Not. R. Astron. Soc.*, 405:525–534, June 2010. doi: 10.1111/j.1365-2966.2010.16466.x.

[59] V. Guillet, A. P. Jones, and G. Pineau Des Forêts. Shocks in dense clouds. II. Dust destruction and SiO formation in J shocks. *AAP*, 497:145–153, April 2009. doi: 10.1051/0004-6361/200811115.

[60] S. Habib, A. Pope, Z. Lukić, D. Daniel, P. Fasel, N. Desai, K. Heitmann, C.-H. Hsu, L. Ankeny, G. Mark, S. Bhattacharya, and J. Ahrens. Hybrid petacomputing meets cosmology: The Roadrunner Universe project. *Journal of Physics Conference Series*, 180(1):012019–+, July 2009. doi: 10.1088/1742-6596/180/1/012019.

[61] A. J. S. Hamilton, P. Kumar, E. Lu, and A. Matthews. Reconstructing the primordial spectrum of fluctuations of the universe from the observed nonlinear clustering of galaxies. *ApJL*, 374:L1–L4, June 1991. doi: 10.1086/186057.

[62] A. P. Hearin and A. R. Zentner. The influence of galaxy formation physics on weak lensing tests of general relativity. *JCAP*, 4:32–+, April 2009. doi: 10.1088/1475-7516/2009/04/032.

[63] A. P. Hearin and A. R. Zentner. The influence of galaxy formation physics on weak lensing tests of general relativity. *JCAP*, 4:32–+, April 2009. doi: 10.1088/1475-7516/2009/04/032.

[64] A. P. Hearin, A. R. Zentner, Z. Ma, and D. Huterer. A General Study of the Influence of Catastrophic Photometric Redshift Errors on Cosmology with Cosmic Shear Tomography. *Astrophys. J.*, 720:1351–1369, September 2010. doi: 10.1088/0004-637X/720/2/1351.

[65] A. Heavens. 3D weak lensing. *Mon. Not. R. Astron. Soc.*, 343:1327–1334, August 2003. doi: 10.1046/j.1365-8711.2003.06780.x.

[66] K. Heitmann, P. M. Ricker, M. S. Warren, and S. Habib. Robustness of Cosmological Simulations. I. Large-Scale Structure. *Astrophys. J. Supp. Ser.*, 160:28–58, September 2005. doi: 10.1086/432646.

[67] K. Heitmann, Z. Lukić, P. Fasel, S. Habib, M. S. Warren, M. White, J. Ahrens, L. Ankeny, R. Armstrong, B. O'Shea, P. M. Ricker, V. Springel, J. Stadel, and H. Trac. The cosmic code comparison project. *Computational Science and Discovery*, 1(1):015003–+, October 2008. doi: 10.1088/1749-4699/1/1/015003.

[68] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon. The Coyote Universe I: Precision Determination of the Nonlinear Matter Power Spectrum. *ArXiv e-prints*, December 2008.

[69] K. Heitmann, D. Higdon, M. White, S. Habib, B. J. Williams, and C. Wagner. The Coyote Universe II: Cosmological Models and Precision Emulation of the Nonlinear Matter Power Spectrum. *ArXiv e-prints*, February 2009.

[70] H. Hoekstra and B. Jain. Weak Gravitational Lensing and Its Cosmological Applications. *Annual Review of Nuclear and Particle Science*, 58:99–123, November 2008. doi: 10.1146/annurev.nucl.58.110707.171151.

[71] H. Hoekstra, H. K. C. Yee, and M. D. Gladders. Constraints on $omega_m$ and σ_8 from weak lensing in red-sequence cluster survey fields. *Astrophys. J.*, 577:595–603, October 2002. doi: 10.1086/342120.

[72] W. Hu. Power Spectrum Tomography with Weak Lensing. *Astrophys. J. Lett.*, 522: L21–L24, September 1999. doi: 10.1086/312210.

[73] W. Hu and B. Jain. Joint galaxy-lensing observables and the dark energy. *Phys. Rev. D*, 70(4):043009–+, August 2004. doi: 10.1103/PhysRevD.70.043009.

[74] W. Hu and I. Sawicki. Models of f(R) cosmic acceleration that evade solar system tests. *Phys. Rev. D*, 76(6):064004–+, September 2007. doi: 10.1103/PhysRevD.76.064004.

[75] W. Hu and I. Sawicki. Parametrized post-Friedmann framework for modified gravity. *Phys. Rev. D*, 76(10):104043–+, November 2007. doi: 10.1103/PhysRevD.76.104043.

[76] W. Hu and M. Tegmark. Weak Lensing: Prospects for Measuring Cosmological Parameters. *Astrophys. J. Lett.*, 514:L65–L68, April 1999. doi: 10.1086/311947.

[77] W. Hu, D. Huterer, and K. M. Smith. Supernovae, the Lensed Cosmic Microwave Background, and Dark Energy. *Astrophys. J. Lett.*, 650:L13–L16, October 2006. doi: 10.1086/508746.

[78] D. Huterer. Weak lensing and dark energy. *Phys. Rev. D*, 65(6):063001, March 2002.

[79] D. Huterer and E. V. Linder. Separating dark physics from physical darkness: Minimalist modified gravity versus dark energy. *Phys. Rev. D*, 75(2):023519–+, January 2007. doi: 10.1103/PhysRevD.75.023519.

[80] D. Huterer and M. Takada. Calibrating the nonlinear matter power spectrum: Requirements for future weak lensing surveys. *Astroparticle Physics*, 23:369–376, May 2005. doi: 10.1016/j.astropartphys.2005.02.006.

[81] D. Huterer and M. White. Nulling tomography with weak gravitational lensing. *Phys. Rev. D*, 72(4):043002–+, August 2005. doi: 10.1103/PhysRevD.72.043002.

[82] D. Huterer, M. Takada, G. Bernstein, and B. Jain. Systematic errors in future weak-lensing surveys: requirements and prospects for self-calibration. *Mon. Not. R. Astron. Soc.*, 366:101–114, February 2006. doi: 10.1111/j.1365-2966.2005.09782.x.

[83] O. Ilbert, P. Capak, M. Salvato, H. Aussel, H. J. McCracken, D. B. Sanders, N. Scoville, J. Kartaltepe, S. Arnouts, E. Le Floc'h, B. Mobasher, Y. Taniguchi, F. Lamareille, A. Leauthaud, S. Sasaki, D. Thompson, M. Zamojski, G. Zamorani, S. Bardelli, M. Bolzonella, A. Bongiorno, M. Brusa, K. I. Caputi, C. M. Carollo, T. Contini, R. Cook, G. Coppa, O. Cucciati, S. de la Torre, L. de Ravel, P. Franzetti, B. Garilli, G. Hasinger, A. Iovino, P. Kampczyk, J.-P. Kneib, C. Knobel, K. Kovac, J. F. Le Borgne, V. Le Brun, O. L. Fèvre, S. Lilly, D. Looper, C. Maier, V. Mainieri, Y. Mellier, M. Mignoli, T. Murayama, R. Pellò, Y. Peng, E. Pérez-Montero, A. Renzini, E. Ricciardelli, D. Schiminovich, M. Scodeggio, Y. Shioya, J. Silverman, J. Surace, M. Tanaka, L. Tasca, L. Tresse, D. Vergani, and E. Zucca. Cosmos Photometric Redshifts with 30-Bands for 2-deg^2. *ApJ*, 690:1236–1249, January 2009. doi: 10.1088/0004-637X/690/2/1236.

[84] M. Ishak. Probing decisive answers to dark energy questions from cosmic complementarity and lensing tomography. *Mon. Not. R. Astron. Soc.*, 363:469–478, October 2005. doi: 10.1111/j.1365-2966.2005.09393.x.

[85] R. K. Jain, P. Chingangbam, J.-O. Gong, L. Sriramkumar, and T. Souradeep. Punctuated inflation and the low CMB multipoles. *Journal of Cosmology and Astroparticle Physics*, 1:9–+, January 2009. doi: 10.1088/1475-7516/2009/01/009.

[86] M. Jarvis, G. M. Bernstein, P. Fischer, D. Smith, B. Jain, J. A. Tyson, and D. Wittman. Weak-Lensing Results from the 75 Square Degree Cerro Tololo Inter-American Observatory Survey. *Astron. J.*, 125:1014–1032, March 2003. doi: 10.1086/367799.

[87] M. Jarvis, B. Jain, G. Bernstein, and D. Dolney. Dark energy constraints from the ctio lensing survey. *Astrophys. J.*, 644:71–79, June 2006. doi: 10.1086/503418.

[88] Y.-P. Jing and L.-Z. Fang. An infrared cutoff revealed by the two years of COBE observations of cosmic temperature fluctuations. *Physical Review Letters*, 73:1882–1885, October 1994. doi: 10.1103/PhysRevLett.73.1882.

[89] Y. P. Jing, P. Zhang, W. P. Lin, L. Gao, and V. Springel. The Influence of Baryons on the Clustering of Matter and Weak-Lensing Surveys. *Astrophys. J. Lett.*, 640:L119–L122, April 2006. doi: 10.1086/503547.

[90] S. Joudaki and M. Kaplinghat. Dark Energy and Neutrino Masses from Future Measurements of the Expansion History and Growth of Structure. *ArXiv e-prints*, June 2011.

[91] S. Joudaki, A. Cooray, and D. E. Holz. Weak lensing and dark energy: The impact of dark energy on nonlinear dark matter clustering. *Phys. Rev. D*, 80(2):023003, July 2009. doi: 10.1103/PhysRevD.80.023003.

[92] S. Jouvel, J.-P. Kneib, O. Ilbert, G. Bernstein, S. Arnouts, T. Dahlen, A. Ealet, B. Milliard, H. Aussel, P. Capak, A. Koekemoer, V. Le Brun, H. McCracken, M. Salvato, and N. Scoville. Designing future dark energy space missions. I. Building realistic

galaxy spectro-photometric catalogs and their first applications. *AAP*, 504:359–371, September 2009. doi: 10.1051/0004-6361/200911798.

[93] G. Jungman, M. Kamionkowski, A. Kosowsky, and D. N. Spergel. Cosmological-parameter determination with microwave background maps. *Phys. Rev. D*, 54:1332–1344, July 1996.

[94] M. Kamionkowski and A. Loeb. Getting around cosmic variance. *Phys. Rev. D*, 56: 4511–4513, October 1997. doi: 10.1103/PhysRevD.56.4511.

[95] N. Katz and S. D. M. White. Hierarchical galaxy formation - Overmerging and the formation of an X-ray cluster. *Astrophys. J.*, 412:455–478, August 1993. doi: 10.1086/172935.

[96] M. Kesden, M. Kamionkowski, and A. Cooray. Can Cosmic Shear Shed Light on Low Cosmic Microwave Background Multipoles? *Physical Review Letters*, 91(22):221302–+, November 2003. doi: 10.1103/PhysRevLett.91.221302.

[97] J. Khoury and A. Weltman. Chameleon Fields: Awaiting Surprises for Tests of Gravity in Space. *Physical Review Letters*, 93(17):171104–+, October 2004. doi: 10.1103/PhysRevLett.93.171104.

[98] J. Khoury and A. Weltman. Chameleon cosmology. *Phys. Rev. D*, 69(4):044026–+, February 2004.

[99] T. D. Kitching, A. F. Heavens, A. N. Taylor, M. L. Brown, K. Meisenheimer, C. Wolf, M. E. Gray, and D. J. Bacon. Cosmological constraints from COMBO-17 using 3D weak lensing. *Mon. Not. R. Astron. Soc.*, 376:771–778, April 2007. doi: 10.1111/j.1365-2966.2007.11473.x.

[100] T. D. Kitching, A. N. Taylor, and A. F. Heavens. Systematic effects on dark energy from 3D weak shear. *Mon. Not. R. Astron. Soc.*, 389:173–190, September 2008. doi: 10.1111/j.1365-2966.2008.13419.x.

[101] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ArXiv e-prints*, March 2008.

[102] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192:18–+, February 2011. doi: 10.1088/0067-0049/192/2/18.

[103] A. Kosowsky, M. Milosavljevic, and R. Jimenez. Efficient cosmological parameter estimation from microwave background anisotropies. *Phys. Rev. D*, 66(6):063007–+, September 2002. doi: 10.1103/PhysRevD.66.063007.

[104] F. Kühnel and D. J. Schwarz. Large-Scale Suppression from Stochastic Inflation. *Physical Review Letters*, 105(21):211302–+, November 2010. doi: 10.1103/PhysRevLett.105.211302.

[105] A. Lasenby and C. Doran. Closed universes, de Sitter space, and inflation. *Phys. Rev. D.*, 71(6):063502–+, March 2005. doi: 10.1103/PhysRevD.71.063502.

[106] E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, and B. Williams. The Coyote Universe. III. Simulation Suite and Precision Emulator for the Nonlinear Matter Power Spectrum. *ApJ*, 713:1322–1331, April 2010. doi: 10.1088/0004-637X/713/2/1322.

[107] G. F. Lewis, A. Babul, N. Katz, T. Quinn, L. Hernquist, and D. H. Weinberg. The Effects of Gasdynamics, Cooling, Star Formation, and Numerical Resolution in Simulations of Cluster Formation. *Astrophys. J.*, 536:623–644, June 2000. doi: 10.1086/308954.

[108] M. Lima and W. Hu. Photometric redshift requirements for self-calibration of cluster dark energy studies. *Phys. Rev. D*, 76(12):123013–+, December 2007. doi: 10.1103/PhysRevD.76.123013.

[109] M. Lima, C. E. Cunha, H. Oyaizu, J. Frieman, H. Lin, and E. S. Sheldon. Estimating the redshift distribution of photometric galaxy samples. *MNRAS*, 390:118–130, October 2008. doi: 10.1111/j.1365-2966.2008.13510.x.

[110] A. Linde. Can we have inflation with Omega greater than 1? *Journal of Cosmology and Astro-Particle Physics*, 5:2–+, May 2003. doi: 10.1088/1475-7516/2003/05/002.

[111] E. V. Linder. Cosmic growth history and expansion history. *Phys. Rev. D*, 72(4):043529–+, August 2005. doi: 10.1103/PhysRevD.72.043529.

[112] E. V. Linder and R. N. Cahn. Parameterized beyond-Einstein growth. *Astroparticle Physics*, 28:481–488, December 2007. doi: 10.1016/j.astropartphys.2007.09.003.

[113] M. Loverde and N. Afshordi. Extended Limber approximation. *Phys. Rev. D*, 78(12):123506–+, December 2008. doi: 10.1103/PhysRevD.78.123506.

[114] Z. Lukić, K. Heitmann, S. Habib, S. Bashinsky, and P. M. Ricker. The Halo Mass Function: High-Redshift Evolution and Universality. *Astrophys. J.*, 671:1160–1181, December 2007. doi: 10.1086/523083.

[115] D. H. Lyth. Large-scale energy-density perturbations and inflation. *Phys. Rev. D*, 31:1792–1798, April 1985.

[116] C.-P. Ma and E. Bertschinger. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *Astrophys. J.*, 455:7–+, December 1995. doi: 10.1086/176550.

[117] C.-P. Ma and J. N. Fry. Deriving the Nonlinear Cosmological Power Spectrum and Bispectrum from Analytic Dark Matter Halo Profiles and Mass Functions. *Astrophys. J.*, 543:503–513, November 2000. doi: 10.1086/317146.

[118] Z. Ma and G. Bernstein. Size of Spectroscopic Calibration Samples for Cosmic Shear Photometric Redshifts. *Astrophys. J.*, 682:39–48, July 2008. doi: 10.1086/588214.

[119] Z. Ma, W. Hu, and D. Huterer. Effects of Photometric Redshift Uncertainties on Weak-Lensing Tomography. *Astrophys. J.*, 636:21–29, January 2006. doi: 10.1086/497068.

[120] A. V. Macciò, A. A. Dutton, F. C. van den Bosch, B. Moore, D. Potter, and J. Stadel. Concentration, spin and shape of dark matter haloes: scatter and the dependence on mass and environment. *Mon. Not. R. Astron. Soc.*, 378:55–71, June 2007. doi: 10.1111/j.1365-2966.2007.11720.x.

[121] R. Mandelbaum, U. Seljak, and C. M. Hirata. A halo mass-concentration relation from weak lensing. *Journal of Cosmology and Astro-Particle Physics*, 8:6–+, August 2008. doi: 10.1088/1475-7516/2008/08/006.

[122] V. E. Margoniner and D. M. Wittman. Photometric Redshifts and Signal-to-Noise Ratios. *ApJ*, 679:31–51, May 2008. doi: 10.1086/528365.

[123] I. G. McCarthy, J. Schaye, T. J. Ponman, R. G. Bower, C. M. Booth, C. Dalla Vecchia, R. A. Crain, V. Springel, T. Theuns, and R. P. C. Wiersma. The case for AGN feedback in galaxy groups. *MNRAS*, 406:822–839, August 2010. doi: 10.1111/j.1365-2966.2010.16750.x.

[124] P. McDonald, H. Trac, and C. Contaldi. Dependence of the non-linear mass power spectrum on the equationof state of dark energy. *MNRAS*, 366:547–556, February 2006. doi: 10.1111/j.1365-2966.2005.09881.x.

[125] R. B. Metcalf. Gravitational lensing of high-redshift Type IA supernovae: a probe of medium-scale structure. *Mon. Not. R. Astron. Soc.*, 305:746–754, May 1999.

[126] M. J. Mortonson and W. Hu. Evidence for horizon-scale power from CMB polarization. *Phys. Rev. D*, 80(2):027301–+, July 2009. doi: 10.1103/PhysRevD.80.027301.

[127] M. J. Mortonson, W. Hu, and D. Huterer. Falsifying Paradigms for Cosmic Acceleration. *ArXiv e-prints*, October 2008.

[128] M. J. Mortonson, C. Dvorkin, H. V. Peiris, and W. Hu. CMB polarization features from inflation versus reionization. *Phys. Rev. D*, 79(10):103519–+, May 2009. doi: 10.1103/PhysRevD.79.103519.

[129] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger. Theory of cosmological perturbations. *Phys. Rep.*, 215:203–333, June 1992. doi: 10.1016/0370-1573(92)90044-Z.

[130] D. Munshi, P. Valageas, L. van Waerbeke, and A. Heavens. Cosmology with weak lensing surveys. *Phys. Rep.*, 462:67–121, June 2008. doi: 10.1016/j.physrep.2008.02.003.

[131] I. Navarro and K. Van Acoleyen. Modified gravity, dark energy and modified Newtonian dynamics. *Journal of Cosmology and Astro-Particle Physics*, 9:6–+, September 2006. doi: 10.1088/1475-7516/2006/09/006.

[132] I. Navarro and K. Van Acoleyen. f(R) actions, cosmic acceleration and local tests of gravity. *Journal of Cosmology and Astro-Particle Physics*, 2:22–+, February 2007. doi: 10.1088/1475-7516/2007/02/022.

[133] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *Astrophys. J.*, 490:493, December 1997. doi: 10.1086/304888.

[134] A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of LCDM Halo Concentrations. *ArXiv e-prints*, 706, June 2007.

[135] J. A. Newman. Calibrating Redshift Distributions beyond Spectroscopic Limits with Cross-Correlations. *Astrophys. J.*, 684:88–101, September 2008. doi: 10.1086/589982.

[136] J. A. Newman, M. C. Cooper, M. Davis, S. M. Faber, A. L. Coil, P. Guhathakurta, D. C. Koo, A. C. Phillips, C. Conroy, A. A. Dutton, D. P. Finkbeiner, B. F. Gerke, D. J. Rosario, B. J. Weiner, C. N. A. Willmer, R. Yan, J. J. Harker, S. A. Kassin, N. P. Konidaris, K. Lai, D. S. Madgwick, K. G. Noeske, G. D. Wirth, A. J. Connolly, N. Kaiser, E. N. Kirby, B. C. Lemaux, L. Lin, J. M. Lotz, G. A. Luppino, C. Marinoni, D. J. Matthews, A. Metevier, and R. P. Schiavon. The DEEP2 Galaxy Redshift Survey: Design, Observations, Data Reduction, and Redshifts. *ArXiv e-prints*, March 2012.

[137] A. J. Nishizawa, M. Takada, T. Hamana, and H. Furusawa. The impact of photometric redshift errors on weak lensing tomography: a clipping method of the catastrophic errors. *ArXiv:1002.2476*, February 2010.

[138] J. Noreña, L. Verde, R. Jimenez, C. Pena-Garay, and C. Gomez. Cancelling out systematic uncertainties. *ArXiv e-prints*, July 2011.

[139] H. Oyaizu, M. Lima, C. E. Cunha, H. Lin, and J. Frieman. Photometric Redshift Error Estimators. *ApJ*, 689:709–720, December 2008. doi: 10.1086/592591.

[140] H. Oyaizu, M. Lima, and W. Hu. Non-linear evolution of f(R) cosmologies II: power spectrum. *ArXiv e-prints*, July 2008.

[141] J. A. Peacock and S. J. Dodds. Non-linear evolution of cosmological power spectra. *Mon. Not. R. Astron. Soc.*, 280:L19–L26, June 1996.

[142] F. R. Pearce, P. A. Thomas, H. M. P. Couchman, and A. C. Edge. The effect of radiative cooling on the X-ray properties of galaxy clusters. *Mon. Not. R. Astron. Soc.*, 317:1029–1040, October 2000.

[143] U.-L. Pen, T. Zhang, L. van Waerbeke, Y. Mellier, P. Zhang, and J. Dubinski. Detection of Dark Matter Skewness in the VIRMOS-DESCART Survey: Implications for Ω_0. *Astrophys. J.*, 592:664–673, August 2003. doi: 10.1086/375734.

[144] S. Perlmutter, G. Aldering, G. Goldhaber, and et al. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *Astrophys. J.*, 517:565–586, June 1999. doi: 10.1086/307221.

[145] Y.-S. Piao, B. Feng, and X. Zhang. Suppressing the CMB quadrupole with a bounce from the contracting phase to inflation. *Phys. Rev. D.*, 69(10):103520–+, May 2004. doi: 10.1103/PhysRevD.69.103520.

[146] A. Pontzen and H. V. Peiris. The cut-sky cosmic microwave background is not anomalous. *Phys. Rev. D.*, 81(10):103008–+, May 2010. doi: 10.1103/PhysRevD.81.103008.

[147] Y. Rasera, J.-M. Alimi, J. Courtin, F. Roy, P.-S. Corasaniti, A. Füzfa, and V. Boucher. Introducing the Dark Energy Universe Simulation Series (DEUSS). In J.-M. Alimi & A. Fuözfa, editor, *American Institute of Physics Conference Series*, volume 1241 of *American Institute of Physics Conference Series*, pages 1134–1139, June 2010. doi: 10.1063/1.3462610.

[148] A. Refregier. Weak Gravitational Lensing by Large-Scale Structure. *Ann. Rev. Astron. Astrophys.*, 41:645–668, 2003. doi: 10.1146/annurev.astro.41.111302.102207.

[149] A. Refregier, R. Massey, J. Rhodes, R. Ellis, J. Albert, D. Bacon, G. Bernstein, T. McKay, and S. Perlmutter. Weak Lensing from Space. III. Cosmological Parameters. *Astron. J.*, 127:3102–3114, June 2004. doi: 10.1086/420986.

[150] A. Refregier, A. Amara, T. D. Kitching, A. Rassat, R. Scaramella, J. Weller, and f. t. Euclid Imaging Consortium. Euclid Imaging Consortium Science Book. *ArXiv:1001.0061*, January 2010.

[151] A. Refregier, A. Amara, T. D. Kitching, A. Rassat, R. Scaramella, J. Weller, and f. t. Euclid Imaging Consortium. Euclid Imaging Consortium Science Book. *ArXiv e-prints*, January 2010.

[152] A. G. Riess, A. V. Filippenko, P. Challis, and et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astron. J.*, 116:1009–1038, September 1998. doi: 10.1086/300499.

[153] A. G. Riess, L.-G. Strolger, J. Tonry, S. Casertano, H. C. Ferguson, B. Mobasher, P. Challis, A. V. Filippenko, S. Jha, W. Li, R. Chornock, R. P. Kirshner, B. Leibundgut, M. Dickinson, M. Livio, M. Giavalisco, C. C. Steidel, T. Benítez, and Z. Tsvetanov.

Type Ia Supernova Discoveries at z ¿ 1 from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. *Astrophys. J.*, 607: 665–687, June 2004. doi: 10.1086/383612.

[154] B. Robertson, A. Kravtsov, J. Tinker, and A. Zentner. Collapse Barriers and Halo Abundance: Testing the Excursion Set Ansatz. *ArXiv e-prints*, December 2008.

[155] D. H. Rudd, A. R. Zentner, and A. V. Kravtsov. Effects of Baryons and Dissipation on the Matter Power Spectrum. *Astrophys. J.*, 672:19, March 2007.

[156] J. Schaye, C. Dalla Vecchia, C. M. Booth, R. P. C. Wiersma, T. Theuns, M. R. Haas, S. Bertone, A. R. Duffy, I. G. McCarthy, and F. van de Voort. The physics driving the cosmic star formation history. *MNRAS*, 402:1536–1560, March 2010. doi: 10.1111/j.1365-2966.2009.16029.x.

[157] R. J. Scherrer and E. Bertschinger. Statistics of primordial density perturbations from discrete seed masses. *Astrophys. J.*, 381:349–360, November 1991. doi: 10.1086/170658.

[158] D. J. Schlegel, C. Bebek, H. Heetderks, S. Ho, M. Lampton, M. Levi, N. Mostek, N. Padmanabhan, S. Perlmutter, N. Roe, M. Sholl, G. Smoot, M. White, A. Dey, T. Abraham, B. Jannuzi, D. Joyce, M. Liang, M. Merrill, K. Olsen, and S. Salim. BigBOSS: The Ground-Based Stage IV Dark Energy Experiment. *ArXiv e-prints*, April 2009.

[159] F. Schmidt. Weak lensing probes of modified gravity. *Phys. Rev. D*, 78(4):043002–+, August 2008. doi: 10.1103/PhysRevD.78.043002.

[160] F. Schmidt, E. Rozo, S. Dodelson, L. Hui, and E. Sheldon. Lensing Bias in Cosmic Shear. *ApJ*, 702:593–602, September 2009. doi: 10.1088/0004-637X/702/1/593.

[161] A. E. Schulz. Calibrating photometric redshift distributions with cross-correlations. *ArXiv:0910.3683*, October 2009.

[162] R. Scoccimarro, R.K̃. Sheth, L. Hui, and B. Jain. How many galaxies fit in a halo? constraints on galaxy formation efficiency from spatial clustering. *Astrophys. J.*, 546: 20–34, January 2001.

[163] U. Seljak. Measuring Polarization in the Cosmic Microwave Background. *ApJ*, 482: 6–+, June 1997. doi: 10.1086/304123.

[164] U. Seljak. Analytic model for galaxy and dark matter clustering. *Mon. Not. R. Astron. Soc.*, 318:203–213, October 2000.

[165] E. Semboloni, L. van Waerbeke, C. Heymans, T. Hamana, S. Colombi, M. White, and Y. Mellier. Cosmic variance of weak lensing surveys in the non-Gaussian regime. *Mon. Not. R. Astron. Soc.*, 375:L6–L10, February 2007. doi: 10.1111/j.1745-3933.2006.00266.x.

[166] E. Semboloni, H. Hoekstra, J. Schaye, M. P. van Daalen, and I. G. McCarthy. Quantifying the effect of baryon physics on weak lensing tomography. *Mon. Not. R. Astron. Soc.*, pages 1461–+, September 2011. doi: 10.1111/j.1365-2966.2011.19385.x.

[167] N. Seto and E. Pierpaoli. Probing the Largest Scale Structure in the Universe with Polarization Map of Galaxy Clusters. *Physical Review Letters*, 95(10):101302, September 2005. doi: 10.1103/PhysRevLett.95.101302.

[168] C. Shapiro and S. Dodelson. Combining weak lensing tomography with halo clustering to probe dark energy. *Phys. Rev. D*, 76(8):083515, October 2007. doi: 10.1103/PhysRevD.76.083515.

[169] R. K. Sheth and G. Tormen. Large-scale bias and the peak background split. *Mon. Not. R. Astron. Soc.*, 308:119–126, September 1999.

[170] R. Sinha and T. Souradeep. Post-WMAP assessment of infrared cutoff in the primordial spectrum from inflation. *Phys. Rev. D.*, 74(4):043518–+, August 2006. doi: 10.1103/PhysRevD.74.043518.

[171] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman. Stable clustering, the halo model and non-linear cosmological power spectra. *Mon. Not. R. Astron. Soc.*, 341:1311–1332, June 2003. doi: 10.1046/j.1365-8711.2003.06503.x.

[172] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. *Astrophys. J.*, 396:L1–L5, September 1992. doi: 10.1086/186504.

[173] Y.-S. Song and L. Knox. Determination of cosmological parameters from cosmic shear data. *Phys. Rev. D*, 70(6):063510, September 2004. doi: 10.1103/PhysRevD.70.063510.

[174] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. *Astrophys. J. Supp. Ser.*, 170:377–408, June 2007. doi: 10.1086/513700.

[175] R. Stanek, D. Rudd, and A. E. Evrard. The Effect of Gas Physics on the Halo Mass Function. *ArXiv e-prints*, September 2008.

[176] R. Stanek, D. Rudd, and A. E. Evrard. The effect of gas physics on the halo mass function. *MNRAS*, 394:L11–L15, March 2009. doi: 10.1111/j.1745-3933.2008.00597.x.

[177] T. Suginohara and J. P. Ostriker. The Effect of Cooling on the Density Profile of Hot Gas in Clusters of Galaxies: Is Additional Physics Needed? *Astrophys. J.*, 507:16–23, November 1998. doi: 10.1086/306326.

[178] L. Sun, Z.-H. Fan, C. Tao, J.-P. Kneib, S. Jouvel, and A. Tilquin. Catastrophic Photo-z Errors and the Dark Energy Parameter Estimates with Cosmic Shear. *Astrophys. J.*, 699:958–967, July 2009. doi: 10.1088/0004-637X/699/2/958.

[179] M. Takada and B. Jain. Cosmological parameters from lensing power spectrum and bispectrum tomography. *Mon. Not. R. Astron. Soc.*, 348:897–915, March 2004. doi: 10.1111/j.1365-2966.2004.07410.x.

[180] M. Takada and M. White. Tomography of Lensing Cross-Power Spectra. *Astrophys. J. Lett.*, 601:L1–L4, January 2004. doi: 10.1086/381870.

[181] M. Tegmark, A. N. Taylor, and A. F. Heavens. Karhunen-Loeve Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets? *ApJ*, 480:22–+, May 1997. doi: 10.1086/303939.

[182] M. Tegmark, M. A. Strauss, M. R. Blanton, and et al. Cosmological parameters from SDSS and WMAP. *Phys. Rev. D*, 69(10):103501, May 2004. doi: 10.1103/PhysRevD.69.103501.

[183] M. Tegmark, D. J. Eisenstein, M. A. Strauss, and et al. Cosmological constraints from the SDSS luminous red galaxies. *Phys. Rev. D*, 74(12):123507, December 2006. doi: 10.1103/PhysRevD.74.123507.

[184] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *Astrophys. J.*, 688:709–728, December 2008. doi: 10.1086/591439.

[185] C. Vale and M. White. Simulating Weak Lensing by Large-Scale Structure. *Astrophys. J.*, 592:699–709, August 2003. doi: 10.1086/375867.

[186] M. P. van Daalen, J. Schaye, C. M. Booth, and C. Dalla Vecchia. The effects of galaxy formation on the matter power spectrum: a challenge for precision cosmology. *MNRAS*, 415:3649–3665, August 2011. doi: 10.1111/j.1365-2966.2011.18981.x.

[187] L. Van Waerbeke, Y. Mellier, and H. Hoekstra. Dealing with systematics in cosmic shear studies: New results from the VIRMOS-Descart survey. *Astron. Astrophys.*, 429: 75–84, January 2005. doi: 10.1051/0004-6361:20041513.

[188] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S. S. Murray, and L. Van Speybroeck. Chandra Sample of Nearby Relaxed Galaxy Clusters: Mass, Gas Fraction, and Mass-Temperature Relation. *Astrophys. J.*, 640:691–709, April 2006. doi: 10.1086/500288.

[189] S. Wang, L. Hui, M. May, and Z. Haiman. Is modified gravity required by observations? An empirical consistency test of dark energy models. *Phys. Rev. D*, 76(6):063503–+, September 2007. doi: 10.1103/PhysRevD.76.063503.

[190] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo. Observational Probes of Cosmic Acceleration. *ArXiv e-prints*, January 2012.

[191] M. White. Baryons and weak lensing power spectra. *Astroparticle Physics*, 22:211–217, November 2004. doi: 10.1016/j.astropartphys.2004.06.001.

[192] M. White and W. Hu. A New Algorithm for Computing Statistics of Weak Lensing by Large-Scale Structure. *Astrophys. J*, 537:1–11, July 2000. doi: 10.1086/309009.

[193] W. M. Wood-Vasey, G. Miknaitis, C. W. Stubbs, and et al. Observational Constraints on the Nature of the Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey. *Astrophys. J. submitted (astro-ph/0701041)*, January 2007.

[194] J. Yokoyama. Chaotic new inflation and primordial spectrum of adiabatic fluctuations. *Phys. Rev. D.*, 59(10):107303–+, May 1999. doi: 10.1103/PhysRevD.59.107303.

[195] A. R. Zentner and S. Bhattacharya. Utilizing Type Ia Supernovae in a Large, Fast, Imaging Survey to Constrain Dark Energy. *ArXiv e-prints*, December 2008.

[196] A. R. Zentner and S. Bhattacharya. Utilizing Type Ia Supernovae in a Large, Fast, Imaging Survey to Constrain Dark Energy. *Astrophys. J.*, 693:1543–1553, March 2009. doi: 10.1088/0004-637X/693/2/1543.

[197] A. R. Zentner, D. H. Rudd, and W. Hu. Self-calibration of tomographic weak lensing for the physics of baryons to constrain dark energy. *Phys. Rev. D*, 77(4):043507–+, February 2008. doi: 10.1103/PhysRevD.77.043507.

[198] H. Zhan. Cosmic tomographies: baryon acoustic oscillations and weak lensing. *Journal of Cosmology and Astro-Particle Physics*, 8:8, August 2006. doi: 10.1088/1475-7516/2006/08/008.

[199] H. Zhan and L. Knox. Effect of Hot Baryons on the Weak-Lensing Shear Power Spectrum. *Astrophys. J. Lett.*, 616:L75–L78, December 2004. doi: 10.1086/426712.

[200] H. Zhan and L. Knox. How Tomographic Cosmic Shear Maps Lead to Constraints on Dark Energy Properties. *ArXiv Astrophysics e-prints*, November 2006.

[201] H. Zhan, L. Knox, and J. A. Tyson. Distance, Growth Factor, and Dark Energy Constraints from Photometric Baryon Acoustic Oscillation and Weak Lensing Measurements. *ArXiv e-prints*, June 2008.

[202] Jun Zhang, Lam Hui, and Albert Stebbins. Isolating geometry in weak lensing measurements. *Astrophys. J.*, 635:806–820, 2005.

[203] P. Zhang, U.-L. Pen, and G. Bernstein. Self calibration of photometric redshift scatter in weak lensing surveys. *ArXiv:0910.4181*, October 2009.

[204] G.-B. Zhao, L. Pogosian, A. Silvestri, and J. Zylberberg. Cosmological tests of GR – a look at the principals. *ArXiv:0905.1326*, May 2009.

[205] G.-B. Zhao, H. Zhan, L. Wang, Z. Fan, and X. Zhang. Probing Dark Energy with the Kunlun Dark Universe Survey Telescope. *PASP*, 123:725–734, June 2011. doi: 10.1086/660801.